Stabilized method of imposing Dirichlet boundary conditions using a recovered stress field

Manuel Tur[1*], José Albelda[1], Onofre Marco[1] y Juan José Ródenas[1]

1: Centro de Investigación en Ingeniería Mecánica
Departamento de Ingeniería Mecánica y de Materiales
Universitat Poliécnica de Valencia
Camino de Vera s/n 46022 Valencia
e-mail:
manuel.tur@mcm.upv.es,jalbelda@mcm.upv.es,onmaral@upvnet.upv.es,jjrodena@mcm.upv.es

Abstract

This paper proposes a new formulation to impose Dirichlet boundary conditions on immersed boundary Cartesian Finite Element meshes. The method uses a recovered stress field calculated by Superconvergent Patch Recovery to stabilize the Lagrange multiplier formulation of the problem. The optimal convergence of the method and the convergence of the proposed iterative procedure are demonstrated. The proposed method is also suitable for problems with non-linear material behavior. Some numerical examples are included to confirm the theoretical results.

Keywords: Dirichlet boundary conditions, Lagrange multipliers, Stabilization, Immersed boundary method, Cartesian grid

1. Introduction

The finite element method (FEM) makes use of a mesh of elements to perform the analysis that will provide the numerical solution to the problem under consideration. In the standard version of the FEM the mesh must conform to the geometry of the domain to be analyzed. At the same time, the distortion of each of the elements with respect to the reference elements in the normalized space must be kept sufficiently low, as high element distortion leads to inaccurate results. As a result, it is no simple task to create an appropriate mesh for finite element analysis (FEA). According to [1], a study at Sandia National Laboratories (USA) revealed that the generation

of the finite element (FE) numerical model, including the process of creating a geometry suitable for analysis by the FEM and the subsequent geometry meshing, takes 80% of the total analysis time, whereas only 20% is devoted to the numerical analysis itself, which provides the solution to the problem. Under the umbrella term of *finite elements in ambient space* [2] we can classify a number of variants of the standard FEM which have recently gained in popularity because they reduce the computational cost to generate the FE model by making the mesh independent of the geometry of the problem. These techniques have been given several names in the literature, such as *Fictitious Domain*, *Implicit Meshing*, *Immersed FEM*, *Immersed Boundary Method*, *Fixed grid FEM* and *Cartesian grid FEM* (cgFEM). In these techniques an auxiliary domain Ω_E with a simple geometry that embeds the problem domain Ω is used for the FE discretization. Because of the simple geometry used to define Ω_E (normally a square in 2D or a cube for the 3D case) its subdivision into elements is very simple, thus reducing the meshing burden. There are two main issues that clearly distinguish these methods from the standard FEM: integration and imposition of boundary conditions, in particular the Dirichlet boundary conditions.

Integration: As the mesh does not conform to the geometry of the domain it is necessary to use special procedures to evaluate integrals. Different approaches have been considered in the bibliography to ensure that the integration in each element is only extended to the exact part of the volume (area in 2D), see for example references [3, 4, 5, 6]. In general terms the solution to this problem consists of using two different meshes, one for interpolation and another for integration. In order to maintain the optimal convergence rate of the FE solution, the degree of approximation to the boundary must be at least of the same order as the degree of the FE interpolation [7]. Transfinite mapping techniques commonly used in the *p*-version of the FEM, or the integration techniques described in [8] to consider the exact geometry given by a NURBS representation of the boundary, can be used in the elements cut by the boundary to obtain an exact representation of the domain.

Boundary conditions: as the FE nodes do not generally lie on the boundary, the procedures used in the standard FEM to apply the boundary conditions cannot be used. The case of the Neumann boundary conditions can be easily tackled by simply taking into account that the integration surface can cut the element and does not necessarily have to coincide with the element faces. However, the case of the Dirichlet boundary conditions is much more complex. To solve the problem, a common alternative is to use the Lagrange

2

multiplier technique. It can be difficult to find compatible discretizations of displacements and multipliers that satisfy the *InfSup* condition [9] but 2D [10] and 3D [11] methods of doing so can be found in the bibliography. However, the naive choice for the multiplier interpolation based on the element edge intersection does not satisfy the *InfSup* condition. The main problem appears because the number of Lagrange multipliers is too high, which can cause undesired oscillations in the Lagrange multiplier field. One of the alternatives is thus to stabilize the solution [12, 13].

One of the most popular methods to stabilize the solution is by Nitche's method, which can be derived from the Barbosa-Hughes stabilization [12]. For early implementations of the method in immersed boundaries see for example [14, 15, 16], or for a comparison with other methods see [11, 17]. The stabilizing term in Nitsche's method has an algorithmic constant to be defined that affects the stability of the method. As pointed out in the bibliography [11, 18, 19], this constant depends on how the boundary intersects the underlying mesh and it can become unbounded for certain configurations. At the same time very large values of the penalty constant result in an ill-conditioned system [20]. This issue has been treated for interface and X-FEM problems by choosing appropriate values for each intersected element [21, 22, 23, 24]. In these works the stabilizing term is the finite element stress field, like in the Barbosa-Hughes stabilization, but computed in a different way. In [21] the stress is computed in the element partition with larger size. In the weighted Nitsche's method [22, 24] the stress field of both partitions are weighted using the partition sizes.

Some variations of Nitsche's method have been proposed in recent years to overcome this problem for imposing Dirichlet constraints. In [18] the solution of the internal elements is extended to the boundary elements with a very small volume/area ratio. The same idea was exploited in [25]. In [19] and [26] the flux jump across the boundary element edges is used to modify the stabilized problem. In [27] the solution of a coarser mesh was used as stabilizing stress. Other stabilizing techniques not directly derived from the Nitsche's method have also been proposed. In [28] the discontinuous-Galerkin method was used. In [29] a polynomial stabilization first proposed in [30] was applied for solving contact problems. Similar ideas where used in [31] where the stabilizing term is a suitable projection of the Lagrange multiplier field. In the method proposed in [32] the Lagrange multipliers field is defined in all the domain of the boundary elements and an optimal value of the penalty parameter is proposed regardless of mesh size.

The aim of this work is to propose a stabilized formulation using a recovered stress field (SPR-type recovered solution, Superconvergent Patch Recovery [33]) iteratively obtained from the finite element solution. We present theoretical values for the stabilizing constant, for linear and quadratic discretizations, that ensure optimal convergence rate. The theoretical analyses are done for the linear elastic case, although it will be also shown that the proposed method can be used to impose Dirichlet boundary conditions in problems with non-linear material behavior, such as plasticity. The paper is organized as follows. Section 2 presents the linear elasticity problem and its mixed finite element implementation. Section 3 presents the stabilized formulation in the context of Cartesian grids. Section 4 describes the finite element interpolation. Section 5 describes the adaptation of the SPR technique required by the proposed method and the evaluation of the stabilization term. Section 6 demonstrates the convergence and stability of the method. The numerical results are given in Section 7, followed by the main conclusions of this work.

2. Statement of the problem

Let us consider the linear elastic problem. Let $\Omega \in \mathbb{R}^d$, with $d = 2$ or $d = 3$ be a bounded domain with a sufficiently smooth boundary Γ. The contour can be divided into two non-overlapping parts, Γ_D and Γ_N, where Dirichlet and Neumann conditions are respectively imposed. The aim is to find the displacement field $\mathbf{u} \in \mathscr{U}$ that fulfills internal equilibrium equations in the domain and the Dirichlet and Neumann boundary conditions on the boundary, which can be written as follows:

$$
\begin{aligned}
\nabla \boldsymbol{\sigma}\left(\mathbf{u}\right) + \mathbf{t}_v = 0 & \quad \text{in} \quad \Omega \\
\boldsymbol{\sigma}\left(\mathbf{u}\right)\mathbf{n} = \mathbf{t}_s & \quad \text{on} \quad \Gamma_\mathrm{N} \\
\mathbf{u} = \mathbf{g} & \quad \text{on} \quad \Gamma_\mathrm{D} \\
\boldsymbol{\epsilon}\left(\mathbf{u}\right) = \mathscr{D}\,\boldsymbol{\sigma}\left(\mathbf{u}\right) &
\end{aligned}
\tag{1}
$$

In the above expression $\mathbf{t}_v \in [L^2(\Omega)]^d$ are the volume forces, $\mathbf{t}_s \in [L^2(\Omega)]^d$ the tractions imposed on the Neumann boundary, $\mathbf{g}$ the displacements imposed on the Dirichlet boundary and $\mathbf{n}$ the unit normal vector. $\mathscr{U} \equiv [H^1(\Omega)]^d$ is the Hilbert space of functions whose integral of the first derivative over the domain is bounded. In linear elasticity, strain tensor is defined from displacement field by

$$
\boldsymbol{\epsilon}\left(\mathbf{u}\right) = \left(\nabla \mathbf{u} + \nabla^\mathrm{T} \mathbf{u}\right)/2
\tag{2}
$$

The constitutive equation, which relates the strains with the stresses by means of the tensor $\mathscr{D}$, can be expressed using two material dependent constants, the Young's modulus E and the Poisson ratio ν, in the case of isotropic behavior. This relationship can be written as

$$\boldsymbol{\epsilon} = \left(\boldsymbol{\sigma} - \nu\left(\operatorname{tr}\left(\boldsymbol{\sigma}\right)\mathbf{I} - \boldsymbol{\sigma}\right)\right)/E \tag{3}$$

It is straightforward to show the following property concerning the constitutive equation, which will be used below.

Property 1. *The scalar product of the tractions can be bounded by the energy per unit volume with a constant C_E, which depends on the material properties, as*

$$\|\boldsymbol{\sigma}(\mathbf{u})\|^2 \leq C_E\left(\boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{u})\right) \qquad with \qquad C_E = \frac{E}{1 - 2\nu} \tag{4}$$

The weak variational formulation of the elastic problem allows two approaches to imposing the Dirichlet boundary conditions. The most common procedure is to impose a constraint in the space of virtual displacement $\mathscr{V}$, i.e. the virtual displacement is zero on the Dirichlet boundary. The virtual work of the elastic forces is in equilibrium with the virtual work of the external forces applied, as follow

$$a\left(\mathbf{u}, \mathbf{v}\right) = c\left(\mathbf{v}\right) \qquad \forall \mathbf{v} \in \mathscr{V}$$
where
$$a\left(\mathbf{u}, \mathbf{v}\right) = \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\epsilon}\left(\mathbf{v}\right)\ d\Omega \tag{5}$$
$$c\left(\mathbf{v}\right) = \int_\Omega \mathbf{v}\cdot\mathbf{t}_v\ d\Omega + \int_{\Gamma_{\mathrm{N}}} \mathbf{v}\cdot\mathbf{t}_s\ d\Gamma$$

This method is simple to implement and effective in the context of the standard finite element method in which the geometry boundary is properly represented by the mesh. However, for Cartesian meshes this method is not valid, since it is very difficult to get a null field on the Dirichlet boundary if the contour of the geometry does not match with the edge of the elements. For this reason it seems more appropriate to seek another formulation, which involves raising the elastic problem as a minimization with constraints. This means finding the displacement field $\mathbf{u}$ that minimizes the total potential energy, subject to the constraints imposed by Dirichlet conditions. The problem

can be expressed as

$$\min \frac{1}{2} a\left(\mathbf{v}, \mathbf{v}\right) - c\left(\mathbf{v}\right)$$
$$\text{with} \quad \mathbf{v} = \mathbf{g} \text{ in } \Gamma_{\mathrm{D}} \tag{6}$$

One approach to solving this minimization problem is to use the Lagrange multiplier method. In addition to the displacement field, a new field of Lagrange multipliers $\boldsymbol{\lambda}$ associated with the reaction forces is added. Formally, the problem is to find the saddle point $[\mathbf{u}, \boldsymbol{\lambda}] \in \mathscr{U} \times \mathscr{M}$ of the following Lagrangian

$$\mathscr{L}\left(\mathbf{v}, \boldsymbol{\mu}\right) = \frac{1}{2} a\left(\mathbf{v}, \mathbf{v}\right) + b\left(\boldsymbol{\mu}, \mathbf{v} - \mathbf{g}\right) - c\left(\mathbf{v}\right) \tag{7}$$

where the Lagrange multipliers belong to the Hilbert space $\mathscr{M} = [H^{-\frac{1}{2}}(\Gamma_{\mathrm{D}})]^d$ and the following functional is defined

$$b\left(\cdot, \cdot\right) : \mathscr{M} \times \mathscr{U} \to \mathbb{R} \qquad b\left(\boldsymbol{\mu}, \mathbf{u}\right) = \int_{\Gamma_{\mathrm{D}}} \boldsymbol{\mu} \cdot \mathbf{u} \, d\Gamma \tag{8}$$

2.1. Finite element formulation

The domain is subdivided into finite elements by a Cartesian mesh in which the boundary of the domain does not necessarily coincide with the edge of an element, but can pass through it. The spaces of the finite element solution are denoted as $\mathscr{U}^h \subset \mathscr{U}$ for displacements and $\mathscr{M}^h \subset \mathscr{M}$ for multipliers. Substituting the finite element fields in Equation (7) and optimizing the Lagrangian we find the following system:

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) + b\left(\boldsymbol{\lambda}^h, \mathbf{v}^h\right) = c\left(\mathbf{v}^h\right) \qquad \forall \mathbf{v}^h \in \mathscr{U}^h$$
$$b\left(\boldsymbol{\mu}^h, \mathbf{u}^h\right) = b\left(\boldsymbol{\mu}^h, \mathbf{g}\right) \qquad \forall \boldsymbol{\mu}^h \in \mathscr{M}^h \tag{9}$$

where $\mathbf{v}^h$ and $\boldsymbol{\mu}^h$ are the variations of the displacement and multiplier fields and $[\mathbf{u}^h, \boldsymbol{\lambda}^h]$ is the solution.

It is well-known [9, 34] that the finite element field of displacements and Lagrange multipliers must fulfill two conditions (*ElKer* and *InfSup*) to obtain an optimal convergence rate of the solution as the mesh is refined. The first of these, ellipticity of $a\left(\cdot, \cdot\right)$ in the kernel of $b\left(\cdot, \cdot\right)$, is easy to fulfill and ensures that there are enough multipliers to prevent rigid body motions. The second prevents too many multipliers being introduced and it is more difficult to

fulfill in practice. There are examples in the bibliography that satisfy the *InfSup* condition, such as the 'vital vertex' method in 2D [10] and 3D [11] for linear discretizations. However, this method has the drawback of increasing the number of unknowns of the problem. In addition, the coefficient matrix of the system is indefinite, which can increase the computational cost as compared to semi-definite positive systems.

The following norms that will be used throughout the text can be defined: the L^2-norm, the energy norm and a mesh dependent norm whose approximation properties can be found in [35, 36]:

$$
\begin{aligned}
\|\mathbf{u}^h\|_{L^2,\Omega}^2 &= \int_{\Omega^h} \mathbf{u}^h \cdot \mathbf{u}^h \, d\Omega \\
\|\mathbf{u}^h\|_E^2 &= \int_{\Omega^h} \boldsymbol{\sigma}(\mathbf{u}^h) : \boldsymbol{\epsilon}(\mathbf{u}^h) \, d\Omega \\
\|\mathbf{u}^h\|_{\mathscr{U}^h}^2 &= \left|\mathbf{u}^h\right|_{H^1,\Omega}^2 + \sum_e h_e^{-1} \|\mathbf{u}^h\|_{L^2,\Gamma_D^e}^2 \\
\|\boldsymbol{\lambda}^h\|_{\mathscr{M}^h}^2 &= \sum_e h_e \|\boldsymbol{\lambda}^h\|_{L^2,\Gamma_D^e}^2
\end{aligned}
\tag{10}
$$

Note that the L^2-norm can also be defined for the boundary $\|\cdot\|_{L^2,\Gamma}^2$ replacing the integration domain. In the last norms, the summation extends to all elements of the mesh that are intersected by Γ_D and h_e is the size of the element intersected by the contour.

3. Stabilized methods

In practice, the problem with the Lagrange multiplier formulation (Equation (9)) is that most natural choices of the Lagrange multiplier field do not fulfill the *InfSup* condition because they introduce too many constraints. The idea behind stabilized methods is to impose additional conditions on Lagrange multipliers without modifying the problem solution, at least at the limit when the element size approaches zero, in order to have more freedom to choose the Lagrange multiplier field.

3.1. Nitsche's method

Nitsche's method is one of the most widely used of the stabilization methods. It is related to the stabilized formulation proposed by Barbosa and Hughes [12, 13] to circumvent the Babuska-Brezzi condition. Its formulation is based

on using the tractions on the boundary as a stabilization term of the multipliers. The original formulation using Lagrange multipliers can be obtained from the following stabilized Lagrangian:

$$\mathscr{L}_N\left(\mathbf{v}^h, \boldsymbol{\mu}^h\right) = \mathscr{L}\left(\mathbf{v}^h, \boldsymbol{\mu}^h\right) - \frac{1}{2}\frac{h}{k}\int_{\Gamma_\mathrm{D}}\left\|\boldsymbol{\mu}^h + \boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}\right\|^2 d\Gamma \qquad (11)$$

where $\mathscr{L}\left(\mathbf{v}, \boldsymbol{\mu}\right)$ was defined in Equation (7) , h is the element size and k is a positive constant having the same units as the Young modulus. In order to use a dimensionless algorithmic constant we define $k = \kappa C_E$, using the constant defined in Equation (4) depending on the material properties. The saddle point of the Lagrangian (11) is $[\mathbf{u}^h, \boldsymbol{\lambda}^h] \in \mathscr{U}^h \times \mathscr{M}^h$ such that:

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) + b\left(\boldsymbol{\lambda}^h, \mathbf{v}^h\right) - \frac{h}{k}\int_{\Gamma_\mathrm{D}}\left(\boldsymbol{\lambda}^h + \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}\right)\cdot\boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}\ d\Gamma = c\left(\mathbf{v}^h\right)$$
$$b\left(\boldsymbol{\mu}^h, \mathbf{u}^h\right) - \frac{h}{k}\int_{\Gamma_\mathrm{D}}\boldsymbol{\mu}^h\cdot\left(\boldsymbol{\lambda}^h + \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}\right)\ d\Gamma = b\left(\boldsymbol{\mu}^h, \mathbf{g}\right)$$

$$(12)$$

$\forall[\mathbf{v}^h, \boldsymbol{\mu}^h] \in \mathscr{U}^h \times \mathscr{M}^h$. Stenberg [36] shows that, for a suitable choice of the multiplier space in L^2, the Lagrange multiplier field can be eliminated element by element from Equation (12) to obtain the classical formulation of the Nitsche's method:
Find $\mathbf{u}^h \in \mathscr{U}^h$ such that:

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) - b\left(\boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}, \mathbf{v}^h\right) - b\left(\boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}, \mathbf{u}^h\right) + \frac{k}{h}\int_{\Gamma_\mathrm{D}}\mathbf{u}^h\cdot\mathbf{v}^h\ d\Gamma =$$
$$c\left(\mathbf{v}^h\right) + \frac{k}{h}\int_{\Gamma_\mathrm{D}}\mathbf{g}\cdot\mathbf{v}^h\ d\Gamma - b\left(\boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}, \mathbf{g}\right) \qquad \forall\mathbf{v}^h \in \mathscr{U}^h$$

$$(13)$$

Nitsche's method has been widely used in the context of immersed boundary mesh to solve interface problems (see for example [16, 17, 23, 24, 37]). However, the original Nitsche's method has some limitations when imposing Dirichlet boundary conditions, as has been pointed out in the bibliography [18, 11, 20, 19]. The optimal convergence rate of the finite element solution can only be achieved if the norm of the tractions on the contour can be bounded by the energy norm, i.e.

$$\left\|\boldsymbol{\sigma}(\mathbf{v}^h)\mathbf{n}\right\|_{L^2,\Gamma_\mathrm{D}} \leq \frac{C_N}{h_e}\|\mathbf{v}^h\|_E \qquad (14)$$

with a constant C_N independent of element size. In the case of immersed boundary meshes, in general, C_N cannot be bounded as the mesh is refined. To illustrate this problem, Figure 1 shows an element of a 2D mesh cut by the boundary of the problem domain. The shaded part indicates the internal area of the element Ω^e. If the boundary comes closer to the element edge as the mesh is refined, the size of the area Ω^e is reduced faster than the size of the boundary Γ_D^e, and the expression (14) is fulfilled with unbounded values of C_N. High values of C_N increase the condition number of the system [20] and tend to overweight the boundary terms with respect to the domain energy, thus resulting in a finite element solution with a large error [19] (see numerical examples).
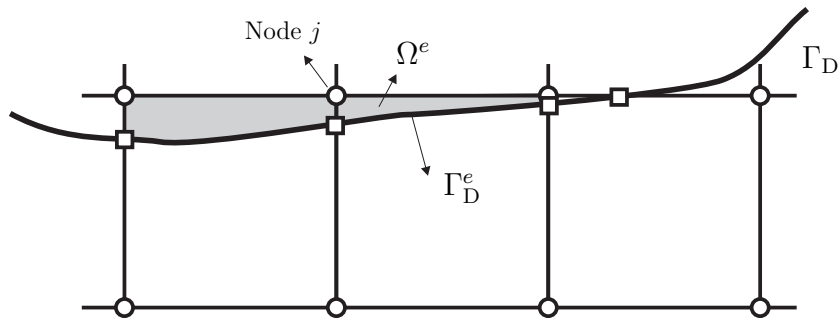

Figure 1: Cartesian element intersected by the geometry with small area.

Remark 1. *In practice, the number of elements with very small internal volume randomly changes between consecutive meshes or in different parts of the boundary, so the effect of this problem is limited. Furthermore, this problem can be minimized by the use of geometric tolerances that avoid the presence of elements with small internal volume, and acceptable results can be obtained from the engineering point of view [11].*

3.2. Proposed method

In this paper we propose the traction computed from a recovered stress field as the stabilization term. The method follows the ideas of extending the solution from the internal elements to the boundary elements [18, 25], but using the stress field instead of the displacement field. The smoothed stress field is obtained from the stresses calculated by the finite element method in the boundary elements and adjacent elements using the concept of the

Superconvergence Patch Recovery [33] (see following section). The aim of this choice is to avoid the problems arising from the condition of Equation (14). As we shall see, this method allows the optimal convergence rate to be obtained for predefined bounded values of the penalty constant k, regardless of the boundary cutting pattern.

The proposed method is derived from the following Lagrangian:

$$\mathscr{L}_S\left(\mathbf{v}^h, \boldsymbol{\mu}^h\right) = \mathscr{L}\left(\mathbf{v}^h, \boldsymbol{\mu}^h\right) - \frac{1}{2}\frac{h}{k}\int_{\Gamma_\mathrm{D}} \left\|\boldsymbol{\mu}^h + \boldsymbol{T}(\hat{\mathbf{u}}^h)\right\|^2 d\Gamma \qquad (15)$$

where $\boldsymbol{T}(\hat{\mathbf{u}}^h)$ is the smoothed traction that depends on the finite element solution computed from a previous iteration, $\hat{\mathbf{u}}^h$. Again the penalty constant can be defined as $k = \kappa C_E$.

The resulting saddle point problem reads as: Find $[\mathbf{u}^h, \boldsymbol{\lambda}^h] \in \mathscr{U}^h \times \mathscr{M}^h$ such that:

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) + b\left(\boldsymbol{\lambda}^h, \mathbf{v}^h\right) = c\left(\mathbf{v}^h\right)$$
$$b\left(\boldsymbol{\mu}^h, \mathbf{u}^h\right) - \frac{h}{k}\int_{\Gamma_\mathrm{D}} \boldsymbol{\mu}^h \cdot \boldsymbol{\lambda}^h \, d\Gamma = b\left(\boldsymbol{\mu}^h, \mathbf{g}\right) + \frac{h}{k}\int_{\Gamma_\mathrm{D}} \boldsymbol{\mu}^h \cdot \boldsymbol{T}(\hat{\mathbf{u}}^h) \, d\Gamma \qquad (16)$$

To obtain Equation (16) from the Lagrangian (15) we considered that $\boldsymbol{T}(\hat{\mathbf{u}}^h)$ is a predefined field, so that its variation is zero. As in Nitsche's method, the proposed formulation can be simplified by eliminating the Lagrange multipliers. Assuming that the multiplier field is in L^2 and following a method proposed in [36] we can solve the second equation in (16) for each element, to obtain the value of the multiplier $\boldsymbol{\lambda}_e^h$ as:

$$\boldsymbol{\lambda}_e^h = \frac{k}{h}\left(\mathbf{u}^h - \mathbf{g}\right) - \boldsymbol{T}(\hat{\mathbf{u}}^h) \qquad (17)$$

Substituting (17) in the first equation of (16) we obtain a modified penalty method: Find the displacement field $\mathbf{u}^h \in \mathscr{U}^h$ such that

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) + \frac{k}{h}\int_{\Gamma_\mathrm{D}} \mathbf{u}^h \cdot \mathbf{v}^h \, d\Gamma =$$
$$c\left(\mathbf{v}^h\right) + \frac{k}{h}\int_{\Gamma_\mathrm{D}} \mathbf{g} \cdot \mathbf{v}^h + \int_{\Gamma_\mathrm{D}} \boldsymbol{T}(\hat{\mathbf{u}}^h) \cdot \mathbf{v}^h \, d\Gamma \qquad \forall \mathbf{v}^h \in \mathscr{U}^h \qquad (18)$$

The second term on the left hand side of (18) is a penalty term with a constant k/h. The last term on the right hand side is the virtual work of

10

reaction forces. It acts as a correction of the penalty term and is necessary for the finite element solution to converge to the exact solution of the problem as the mesh is refined.

As the traction field $\boldsymbol{T}$ depends on the finite element solution, an iterative procedure is proposed to solve the problem. In the first iteration, we solve the problem assuming that $\boldsymbol{T} = 0$. Then the smooth stress field is calculated. This stress field is used to update $\boldsymbol{T}$ in Equation (18) in order to solve the next iteration. This process runs until convergence is achieved.

The advantages of the proposed method are:

- As in Nitsche's method, the unknowns of the problem in the proposed formulation are the displacement degrees of freedom as the multipliers are condensed. Thus the size of the problem is not increased.

- Compared to Nitsche's method (Equation (13)), fewer integral terms are needed to compute the system. As we will see in the following section, the proposed method is stable and convergent in spite of the lacking of these terms.

- The method is stable for a mesh independent bounded value of the penalty constant κ. In the following section we obtain the value of this constant for 8-node linear and 20-node quadratic elements.

- The proposed method can be directly applied to solve problems with non-linear constitutive material behavior (see section 3.4).

The obvious drawback is that multiple iterations are needed to get the solution. However, this disadvantage can be minimized taking into account that the matrix to solve for each iteration is always the same for linear problems (since it is only mesh dependent). Therefore it is only necessary to factorize the matrix once and perform backward substitution every iteration.

3.3. Iterative Nitsche's method

The proposed method can also be used to define another formulation, by replacing the stabilizing smooth stress field $\boldsymbol{T}$ by the traction computed from the finite element solution of a previous iteration $\boldsymbol{\sigma}(\hat{\mathbf{u}}^h)\mathbf{n}$ instead of the recovered tractions. This formulation is also compared in the numerical results.

3.4. Plasticity

Although the equations and all theoretical analysis in this paper are done for the linear elastic case, the proposed method can be used to solve problems with elasto-plastic material behavior. In this case, Equation (18) is still valid if we replace the bilinear form $a\left(\mathbf{u}^h, \mathbf{v}^h\right)$ by the virtual work of internal forces. In the context of immersed boundaries, this term is computed in exactly the same way as is done for the standard finite element method, provided that the quadrature points are properly defined in the internal part of the elements. The stabilizing terms depend on the displacement field and the stabilizing stress $\boldsymbol{T}$. As we shall see, the stress $\boldsymbol{T}$ is directly computed from the quadrature points of the domain (not the boundary) thus there is no need to change anything with respect to the linear elastic case. The method runs in two loops, one to update the stabilizing stress $\boldsymbol{T}$ and another, with $\boldsymbol{T}$ remaining constant, to solve the non-linear plasticity problem. In the numerical examples a plasticity problem is included to show the performance of the method.

4. Interpolation and numerical integration

In this work we use Cartesian meshes formed by 3D hexahedral elements whose contours are parallel to the Cartesian planes and consider linear 8-node $\mathcal{L}_8$ and quadratic 20-node $\mathcal{Q}_{20}$ hexahedral elements. For displacements $\mathbf{u}^h$ the usual finite element interpolation shape functions are defined using degree $p = 1$ for $\mathcal{L}_8$ elements and degree $p = 2$ for $\mathcal{Q}_{20}$ elements.
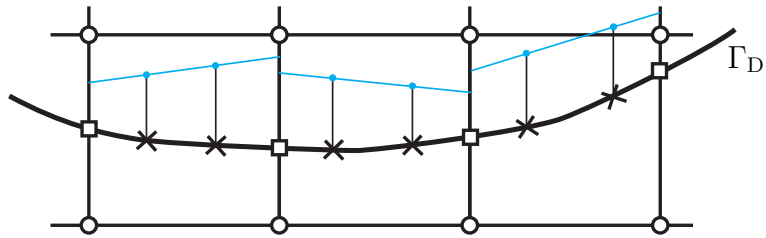

Figure 2: Examples of intersected linear elements in 2D. Segmentation of the boundary Γ_D based on the intersection of the boundary with the element edges (squares). The '×' symbols denote the quadrature points where the Lagrange multipliers are used to define a piecewise discontinuous linear interpolation.

The boundary integrals in the proposed formulation (Equation (18)) are numerically evaluated using Gaussian quadrature. This is equivalent to implicitly define a multiplier unknown at each quadrature point in the saddle point

formulation (Equation (16)) and to eliminate it from the system of equations at element level. This implicit definition was proposed in [38] and used in [27]. A schematic representation of the interpolation is depicted in Figure 2, which shows three 2D elements of the boundary. The quadrature points at the boundary are used to define a polynomial in each boundary segment. A piecewise discontinuous interpolation is then obtained. In practice there is no need to explicitly define this polynomial, because we only need to evaluate it by numerical integration at the quadrature points, where the value is precisely that of the Lagrange multiplier.

In 2D problems, such as the one depicted in Figure 2, n_g is the number of quadrature points at each element. To exactly integrate the product of displacements (degree p) and multipliers (degree $n_g - 1$), if Gaussian quadrature is used, we need to fulfill that $2 \cdot n_g - 1 > p + n_g - 1$, so $n_g > p$. In 3D problems the part of the real boundary in each element is triangulated, and a piecewise discontinuous polynomial interpolation is defined in each triangle (see Figure 3). We use $n_g = 7$ quadrature points, which means that a polynomial of degree 5 can be exactly integrated. The implicit interpolation of the multiplier has a complete degree $q = 2$.
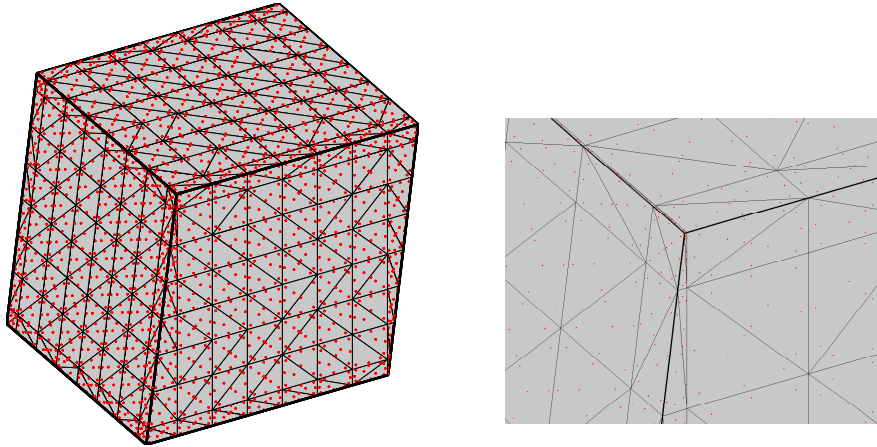

Figure 3: Triangulation of the boundary used to perform the numerical integration. The dots are located at the quadrature points (7 each facet). Complete domain and zoom at a vertex.

The best approximation error of the finite element spaces measured in mesh-dependent norms is discussed in [35, 36]. If the exact solution is smooth enough, namely, if $\mathbf{u} \in H^{p+1}(\Omega)$ and $\boldsymbol{\lambda} \in H^{q+1}(\Gamma)$ and the degree of dis-

placement and multiplier interpolation are at least p and q respectively, it can be proved that the chosen finite element fields fulfill

$$\inf_{\mathbf{v}^h \in \mathscr{U}^h} \|\mathbf{u} - \mathbf{v}^h\|_{\mathscr{U}^h} \le Ch^p \|\mathbf{u}\|_{p+1}$$
$$\inf_{\boldsymbol{\mu}^h \in \mathscr{M}^h} \|\boldsymbol{\lambda} - \boldsymbol{\mu}^h\|_{\mathscr{M}^h} \le Ch^{q+3/2} \|\boldsymbol{\lambda}\|_{q+1,\Gamma} \tag{19}$$

5. Recovered stress field

As mentioned above, the term $\boldsymbol{T}$ used to stabilize the Lagrange multipliers is a recovered stress field obtained from the finite element solution. The construction of the smoothing field is based on the SPR technique [33]. The computation is performed as follows:

1. For each vertex node i of the boundary elements (those intersected by the geometry), including the vertex nodes located outside the domain, we construct a patch with all the elements that contain this node. For illustration purposes, Figure 4 shows part of a 2D mesh close to a boundary of the domain. The elements of the patch are used to calculate the recovered stress at node i. In this case, two elements of the patch are internal and two are divided by the boundary. The shaded region is the part of the domain corresponding to the patch and is denoted as Ω_{patch}.
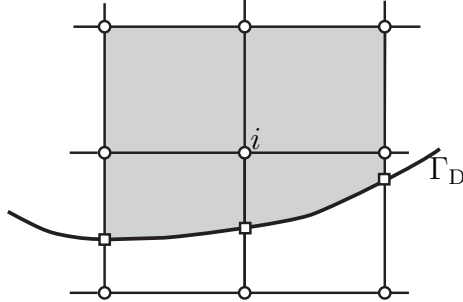

Figure 4: Elements used to compute the smooth stress field for node i.

2. Each component of the recovered stress field for node i is defined as a complete polynomial of degree p (being p the degree of the displacement interpolation) $\boldsymbol{S}_{ic} = \mathbf{x} \cdot \mathbf{w}_{ic}$, where $\mathbf{x} = \{1, x, y, z, xy, \ldots\}$ is the

14

polynomial basis and $\mathbf{w}_{ic}$ is the coefficient vector. Subindex c is introduced to indicate the stress component $c = xx, yy, zz, xy, yz, zx$. These coefficients are calculated by minimizing the L^2-norm of the difference between the polynomial and the finite element solution in the patch, i.e. solving the following minimization problem:

$$
\begin{aligned}
\mathbf{w}_{ic} = \arg\min_{\mathbf{w}} \int_{\Omega_{patch}} (\boldsymbol{\sigma}_c^h - \boldsymbol{S}_{ic}) \cdot (\boldsymbol{\sigma}_c^h - \boldsymbol{S}_{ic})\, d\Omega \approx \\
\arg\min_{\mathbf{w}} \sum_{\forall g \in \Omega_{patch}} H_g J_g (\boldsymbol{\sigma}_g^h - \boldsymbol{S}_g) \cdot (\boldsymbol{\sigma}_g^h - \boldsymbol{S}_g)
\end{aligned}
\tag{20}
$$

The integral is evaluated numerically with the same quadrature points used to calculate the stiffness matrix. The subscript g indicates that the variable is calculated at the quadrature point. H_g is the weight and J_g the Jacobian of the element subtriangulation performed to compute the volume integrals.

3. The recovered stress field in the domain $\boldsymbol{S}_c$ is an interpolation of the nodal polynomials obtained in step 2, $\boldsymbol{S}_c = \sum_{\forall i} N_i \boldsymbol{S}_{ic}$, by using the linear shape functions of the vertex nodes, N_i.

4. The stabilizing stress is the traction computed from the recovered stress tensor: $\boldsymbol{T} = \boldsymbol{S}\mathbf{n}$, where $\mathbf{n}$ is the normal vector and $\boldsymbol{S}$ is the tensor whose components are the $\boldsymbol{S}_c$ polynomials defined above .

The recovered stress polynomial $\boldsymbol{S}_i$ at each patch depends on the solution evaluated at both, boundary elements and internal elements. The integral used to calculate the field $\boldsymbol{S}_i$ gives weight to the integration points as a function of the volume associated to each of them. Therefore, the smaller the volume of the cut elements, the lower the weight and the smaller the influence on the smooth stress field.

The above definition of the recovered field fulfills three properties that will be used in the paper:

Property 2. *The L^2-norm of the recovered stress field, can be bounded by the L^2-norm of the finite element stress field in the solid domain Ω:*

$$
\|\boldsymbol{S}\|_{L^2,\Omega}^2 \leq C_r \left\|\boldsymbol{\sigma}(\mathbf{u}^h)\right\|_{L^2,\Omega}^2
\tag{21}
$$

with $C_r = 8$. Assuming a smooth enough exact solution $C_r \to 1$ as the mesh is refined.

Proof: Taking the derivative in Equation (20) we obtain the following system to evaluate the coefficient vector $\mathbf{w}_i$

$$\left(\int_{\Omega_{patch}} \mathbf{x}^T \mathbf{x} \, d\Omega \right) \mathbf{w}_{ic} = \int_{\Omega_{patch}} \mathbf{x}^T \boldsymbol{\sigma}_c^h \, d\Omega$$

Multiplying the previous expression by the solution $\mathbf{w}_{ic}^T$ we have:

$$\|\boldsymbol{S}_{ic}\|_{L^2,\Omega_{patch}}^2 = \int_{\Omega_{patch}} \boldsymbol{S}_{ic} \cdot \boldsymbol{\sigma}_c^h \, d\Omega$$

Taking into account the last expression and Equation (20) we obtain:

$$\int_{\Omega_{patch}} (\boldsymbol{\sigma}_c^h - \boldsymbol{S}_{ic}) \cdot (\boldsymbol{\sigma}_c^h - \boldsymbol{S}_{ic}) \, d\Omega \geq 0$$

$$\|\boldsymbol{\sigma}_c^h\|_{L^2,\Omega_{patch}}^2 + \|\boldsymbol{S}_{ic}\|_{L^2,\Omega_{patch}}^2 - 2 \int_{\Omega_{patch}} \boldsymbol{\sigma}_c^h \cdot \boldsymbol{S}_{ic} \, d\Omega =$$

$$\|\boldsymbol{\sigma}_c^h\|_{L^2,\Omega_{patch}}^2 - \|\boldsymbol{S}_{ic}\|_{L^2,\Omega_{patch}}^2 \geq 0$$

As each nodal component $\boldsymbol{S}_{ic}$ can be bounded by the stress norm in every patch, the norm of the tensor is also bounded.

On the other hand, the recovered stress field (step 3) is the interpolation of the nodal polynomials using the linear shape functions, whose value is between 0 and 1. Then, using the Cauchy-Schwartz inequality and taking into account that the number of nodes per elements is eight and that the patches contain at most eight elements, we obtain that $C_r = 8$ from:

$$\|\boldsymbol{S}\|_{L^2,\Omega}^2 = \sum_{\forall e} \int_{\Omega^e} \left(\sum_{i=1}^{8} N_i \boldsymbol{S}_i \right)^2 d\Omega \leq \sum_{\forall e} \int_{\Omega^e} \sum_{i=1}^{8} (N_i)^2 \sum_{i=1}^{8} (\boldsymbol{S}_i)^2 \, d\Omega \leq$$

$$\sum_{\forall e} \int_{\Omega^e} \sum_{i=1}^{8} (\boldsymbol{S}_i)^2 \, d\Omega = \sum_{\forall patch} \|\boldsymbol{S}_i\|_{L^2,\Omega_{patch}}^2 \leq \sum_{\forall patch} \|\boldsymbol{\sigma}^h\|_{L^2,\Omega_{patch}}^2 = 8 \|\boldsymbol{\sigma}(\mathbf{u}^h)\|_{L^2,\Omega}^2$$

The value of $C_r = 8$ could be improved by taking into account the definition of the linear shape functions.

Considering that the exact solution is smooth enough, we can assume that the recovered stress field tends to be uniform in each patch (each $\boldsymbol{S}_i$ is constant) as the element size becomes smaller. Under that assumption it follows that:

$$\|\boldsymbol{S}\|_{L^2,\Omega}^2 = \sum_{\forall e} \int_{\Omega^e} \left(\sum_{i=1}^{8} N_i \boldsymbol{S}_i\right)^2 d\Omega \leq \frac{1}{8} \sum_{\forall e} \int_{\Omega^e} \sum_{i=1}^{8} (\boldsymbol{S}_i)^2 \, d\Omega = \left\|\boldsymbol{\sigma}(\mathbf{u}^h)\right\|_{L^2,\Omega}^2$$

and the constant $C_r = 1$.
$\square$

Property 3. *Assuming that every boundary element is connected to at least one internal element, the L^2-norm of $\boldsymbol{T}$ in the boundary can be bounded by the energy norm of the finite element solution with a constant C_p independent of the mesh*

$$\|\boldsymbol{T}\|_{L^2,\Gamma_D}^2 \leq \frac{C_E \, C_r \, C_p}{h} \left\|\mathbf{u}^h\right\|_E^2 \tag{22}$$

where h is the uniform element size, C_E is a material dependent constant defined in (4) and C_r is defined in property 2.
The value of the constant is $C_p = 13$ for $\mathcal{L}_8$ elements and $C_p = 21$ for $\mathcal{Q}_{20}$ elements.

Proof: First we want to bound the L^2-norm of $\boldsymbol{S}$ on the boundary with its norm in the domain. For a given boundary element (e), there exists at least one internal element (ie) that shares an element face. We assume that the worst case to bound Equation (22) occurs when the boundary of the domain practically coincides with an element face. This is schematically depicted in Figure 1 in 2D. Then we have to find the best value of C_p such that the following inequality holds:

$$\|\boldsymbol{S}\|_{L^2,\Gamma^e}^2 = \int_{\Gamma^{ie}} \boldsymbol{S}^2 \, d\Gamma \leq h \, C_p \, \|\boldsymbol{S}\|_{L^2,\Omega^{ie}}^2 = h \, C_p \int_{\Omega^{ie}} \boldsymbol{S}^2 \, d\Omega.$$

$\boldsymbol{S}$ is the product of two polynomials: the smooth field of degree p and the linear shape functions having terms of at most degree 1 in each direction. Thus $\boldsymbol{S}^2$ has terms of degree $2(p+1)$. The above integrals can be evaluated in the reference $[-1, 1]$ element:

$$h^2 \int_{\square,\Gamma} \boldsymbol{S}^2 \, d\Gamma \leq C_p \, h^3 \int_{\square,\Omega} \boldsymbol{S}^2 \, d\Omega$$

17

where the symbol $\Box, \Gamma$ denotes the reference element face and $\Box, \Omega$ its volume. We can use the Newton-Cotes quadrature with positive weights to exactly evaluate the above integrals. The volume integral can be bounded as

$$\int_{\Box,\Omega} \boldsymbol{S}^2 \, d\Omega = H_1 \underbrace{\sum_{j=1}^{n_q} \sum_{k=1}^{n_q} H_j H_k \boldsymbol{S}^2}_{\int_{\Box,\Gamma} \boldsymbol{S}^2 \, d\Gamma} + \underbrace{\sum_{i=2}^{n_q} H_i \sum_{j=1}^{n_q} \sum_{k=1}^{n_q} H_j H_k \boldsymbol{S}^2}_{\text{positive}} \leq H_1 \int_{\Box,\Gamma} \boldsymbol{S}^2 \, d\Gamma$$

where H_1 is the weight of the first quadrature point and n_q the number of quadrature points. For linear elements, $\boldsymbol{T}^2$ is a polynomial of degree 4 and it results that $n_q = 5$ and $C_p = 13$. For quadratic elements the degree of the polynomial is 6, $n_q = 7$ and $C_p = 21$.

Now, it holds that $\|\boldsymbol{T}\|_{L^2,\Gamma^e}^2 \leq \|\boldsymbol{S}\|_{L^2,\Gamma^e}^2$. Adding the contribution of all boundary elements, using Property 2 (Equation (21)) and taking into account the relationship between the L^2 norm of the stresses and the energy norm (Equation (4)) the result follows.

$\Box$

Property 4. *Let $\mathbf{u}$ be the exact solution of the problem. If $\mathbf{u}$ is assumed to be regular enough, the recovered traction evaluated for the exact solution $\boldsymbol{T}(\mathbf{u})$ fulfills the following property:*

$$\|\boldsymbol{\sigma}(\mathbf{u})\boldsymbol{n} - \boldsymbol{T}(\mathbf{u})\|_{L^2,\Gamma_D}^2 \approx \mathcal{O}(h^{p+1}) \tag{23}$$

where h is the element size.

Proof: Let $T(\mathbf{u})$ be the boundary tractions evaluated from (20) by replacing $\boldsymbol{\sigma}^h$ by the exact stress. Therefore, $T(\mathbf{u})$ is a polynomial approximation to the exact traction in each patch. The result follows assuming that $\mathbf{u}$ can be expanded in its Taylor series.

Remark 2. *The SPR technique can be modified to fulfill certain equations, such as equilibrium of stresses, compatibility equations, imposed boundary conditions, etc. This modification improves the approximation of the recovered stress field [7, 39].*

Remark 3. *Although we have chosen the SPR technique, any recovered stress field that satisfies Properties 2, 3 and 4 could be used as the stabilizing term and the results of the following section would hold.*

6. Convergence of the finite element solution

In this section the convergence of the finite element solution is analyzed. We proceed in four steps: First we show that the bilinear form defining the problem is coercive. Then we will analyze the conditions under which the iterative method converges. Finally we will show the stability and convergence to the exact solution of the problem.

6.1. Coercivity

As pointed out above, problem (18) is iteratively solved with an initial value of the stabilization field $\boldsymbol{T} = 0$. The system can be expressed in matrix form as

$$\mathbf{A}\,\mathbf{d}^i = \mathbf{c} + \mathbf{B}\,\mathbf{d}^{i-1} \tag{24}$$

where $\mathbf{d}^i$ is the vector of nodal displacements at iteration i, $\mathbf{A}$ is the matrix on the left side of the Equation (18), $\mathbf{B}$ is the matrix obtained from the last term on the right side of (18) which depends on the method used to calculate the recovered stress field, and $\mathbf{c}$ is the vector derived from the other terms of the right hand side in (18).

In order to check the convergence of the iterative method, we define the residual of the equation as:

$$\mathbf{r}^i = \mathbf{A}\,\mathbf{d}^i - \mathbf{B}\,\mathbf{d}^{i-1} - \mathbf{c} \tag{25}$$

Convergence is achieved when the norm of the residual $\|\mathbf{r}^i\|$ as well as the norm of the difference between two consecutive solutions $\|\mathbf{d}^i - \mathbf{d}^{i-1}\|$ are lower than the given tolerances.

Remark 4. *Problem* (24) *can be solved without explicitly defining matrix* $\boldsymbol{B}$ *by directly computing the product* $\boldsymbol{B}\,\boldsymbol{d}^{i-1}$, *which is easily computed as a surface integral. This term corresponds to the equivalent nodal forces imposed by the stabilization traction.*

The bilinear form defined in Equation (18) from which the matrix $\mathbf{A}$ is obtained is:

$$a\left(\mathbf{u}^h, \mathbf{v}^h\right) + \frac{h}{k}\int_{\Gamma_{\mathrm{D}}} \mathbf{u}^h \cdot \mathbf{v}^h \, d\Gamma$$

For any $k > 0$ the bilinear form is coercive since $a\left(\cdot, \cdot\right)$ is symmetric and positive semi-definite, with $a\left(\mathbf{v}, \mathbf{v}\right) = 0$ only for rigid-body motions. The penalty term ensures that rigid-body motions are not allowed. As a consequence, matrix $\mathbf{A}$ is invertible.

6.2. Convergence of the iterative method

The procedure set out to solve the problem can be considered as a Richardson's iterative method of solving linear systems of equations. This method is known to converge [40] if the spectral radius of the matrix $\mathbf{A}^{-1}\mathbf{B}$ is lower than 1. We obtain the following result:

Proposition 1. *The iterative procedure defined in (24) converges for a large enough but bounded value of the penalty constant $k > C_E C_p C_r / 4$ (or $\kappa > C_p C_r / 4$).*

Proof: The spectral radius of the matrix $\mathbf{A}^{-1}\mathbf{B}$ is defined as the maximum of its eigenvalue modulus. Any eigenvalue λ of this matrix fulfills

$$\mathbf{A}^{-1}\mathbf{B}\,\mathbf{d} = \lambda \mathbf{d} \quad \rightarrow \quad \mathbf{B}\,\mathbf{d} = \lambda \mathbf{A}\,\mathbf{d}$$

Premultiplying by $\mathbf{d}^T$ on both sides of the equation it follows

$$\mathbf{d}^T \mathbf{B}\,\mathbf{d} = \lambda \mathbf{d}^T \mathbf{A}\,\mathbf{d}$$

To prove that the modulus of λ is less than 1, we can see that the left side of the equation corresponds to the stabilization term for a given $\mathbf{v}^h$, so that

$$\mathbf{d}^T \mathbf{B}\,\mathbf{d} = \int_{\Gamma_D} \mathbf{v}^h \cdot \boldsymbol{T}(\mathbf{v}^h)\, d\Gamma \tag{26}$$

Analogously, we can write:

$$\mathbf{d}^T \mathbf{A}\,\mathbf{d} = a\left(\mathbf{v}^h, \mathbf{v}^h\right) + \frac{k}{h} \int_{\Gamma_D} \mathbf{v}^h \cdot \mathbf{v}^h\, d\Gamma \tag{27}$$

Applying the Cauchy-Schwarz inequality to Equation (26), using Equation (22) and considering that for two positive numbers x and y it holds that $2xy \leq x^2 + y^2$, we obtain:

$$\lambda \mathbf{d}^T \mathbf{A}\mathbf{d} = \mathbf{d}^T \mathbf{B}\mathbf{d} \leq \|\mathbf{v}^h\|_{L^2,\Gamma_D} \|\boldsymbol{T}\|_{L^2,\Gamma_D} \leq \|\mathbf{v}^h\|_{L^2,\Gamma_D} \left(\frac{C_E\, C_p\, C_r}{h}\right)^{1/2} \|\mathbf{v}^h\|_E$$

$$\leq \|\mathbf{v}^h\|_E^2 + \frac{C_E\, C_p\, C_r}{4h} \|\mathbf{v}^h\|_{L^2,\Gamma_D}^2 \tag{28}$$

Comparing (27) with (28), it follows that if $k > \dfrac{C_E\, C_p\, C_r}{4}$, the modulus of any eigenvalue λ must be less than 1.

$\square$

20

6.3. Stability of the formulation

We define the following bilinear form associated to our problem (18):

$$Q(\mathbf{w}^h, \mathbf{v}^h) = a\left(\mathbf{w}^h, \mathbf{v}^h\right) + \frac{k}{h} \int_{\Gamma_D} \mathbf{w}^h \cdot \mathbf{v}^h \, d\Gamma - \int_{\Gamma_D} \boldsymbol{T}(\mathbf{w}^h) \cdot \mathbf{v}^h \, d\Gamma \qquad (29)$$

Here we prove the weak coercivity of the functional Q that will be used to demonstrate the optimal convergence of the proposed method. Considering first the last term of Q, we apply the Cauchy-Schwartz inequality and Equation (22) to obtain:

$$-\left| \int_{\Gamma_D} \boldsymbol{T}(\mathbf{v}^h) \cdot \mathbf{v}^h \, d\Gamma \right| \geq -\|\mathbf{v}^h\|_{L^2,\Gamma_D} \|\boldsymbol{T}(\mathbf{v}^h)\|_{L^2,\Gamma_D} \geq$$

$$- \|\mathbf{v}^h\|_{L^2,\Gamma_D} \left(\frac{C_E \, C_p \, C_r}{h_e} \right)^{1/2} \|\mathbf{v}^h\|_E \geq -\frac{1}{2} \|\mathbf{v}^h\|_E^2 - \frac{C_E \, C_p \, C_r}{2h_e} \|\mathbf{v}^h\|_{L^2,\Gamma_D}^2 \tag{30}$$

Proposition 2. *The bilinear form Q is weakly coercive if the penalty parameter is chosen as $k > C_E \, C_p \, C_r$, that is:*

$$\sup_{\mathbf{v}^h \in \mathscr{U}^h} \frac{Q(\mathbf{w}^h, \mathbf{v}^h)}{\|\mathbf{v}^h\|_{\mathscr{U}^h}} \geq \beta \|\mathbf{w}^h\|_{\mathscr{U}^h} \qquad \forall \mathbf{w}^h \in \mathscr{U}^h \tag{31}$$

with $\beta = \dfrac{1}{2}\left(k - C_E \, C_p \, C_r\right)$.

Proof: It suffices to show that the inequality holds for a certain value of $\mathbf{v}^h = \mathbf{w}^h$. Using (30) and the definition of Q (Equation (36)), and considering the mesh dependent norm defined in (10), we have:

$$Q(\mathbf{v}^h, \mathbf{v}^h) = a\left(\mathbf{v}^h, \mathbf{v}^h\right) + \frac{k}{h} \int_{\Gamma_D} \mathbf{v}^h \cdot \mathbf{v}^h \, d\Gamma - \int_{\Gamma_D} \boldsymbol{T}(\mathbf{v}^h) \cdot \mathbf{v}^h \, d\Gamma \geq$$

$$a\left(\mathbf{v}^h, \mathbf{v}^h\right) + \frac{k}{h} \int_{\Gamma_D} \mathbf{v}^h \cdot \mathbf{v}^h \, d\Gamma - \frac{1}{2} \|\mathbf{v}^h\|_E^2 - \frac{C_E \, C_p \, C_r}{2h_e} \|\mathbf{v}^h\|_{L^2,\Gamma_D}^2 \geq$$

$$\frac{1}{2}\left(k - C_E \, C_p \, C_r\right) \|\mathbf{v}^h\|_{\mathscr{U}^h}^2 \tag{32}$$

21

6.4. Optimal convergence

Let $[\mathbf{u}, \boldsymbol{\lambda}]$ be the exact solution of the problem (7) and $[\mathbf{u}^h, \boldsymbol{\lambda}^h]$ the solution of the discretized problem (16). Taking variations in (7), a system is obtained that must be fulfilled $\forall [\mathbf{v}, \boldsymbol{\mu}] \in \mathscr{U} \times \mathscr{M}$. In particular, it will also be true for $\mathbf{v}^h \in \mathscr{U}^h \subset \mathscr{U}$ and $\boldsymbol{\mu}^h \in \mathscr{M}^h \subset \mathscr{M}$, and we can write:

$$
\begin{aligned}
a\left(\mathbf{u}, \mathbf{v}^h\right) + b\left(\boldsymbol{\lambda}, \mathbf{v}^h\right) &= c\left(\mathbf{v}^h\right) \\
b\left(\boldsymbol{\mu}^h, \mathbf{u}\right) &= b\left(\boldsymbol{\mu}^h, \mathbf{g}\right)
\end{aligned}
\tag{33}
$$

Adding the stabilization term $\dfrac{h}{k} \int_{\Gamma_\mathrm{D}} \boldsymbol{\mu}^h \cdot (\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} + \boldsymbol{\lambda}) \ d\Gamma$, (which is zero since for the exact solution $\boldsymbol{\lambda} = -\boldsymbol{\sigma}(\mathbf{u})\mathbf{n}$), to the second equation and subtracting (16) it follows that:

$$
a\left(\mathbf{u} - \mathbf{u}^h, \mathbf{v}^h\right) + b\left(\boldsymbol{\lambda} - \boldsymbol{\lambda}^h, \mathbf{v}^h\right) = 0
$$
$$
b\left(\boldsymbol{\mu}^h, \mathbf{u} - \mathbf{u}^h\right) - \frac{h}{k} \int_{\Gamma_\mathrm{D}} \boldsymbol{\mu}^h \cdot \left(\boldsymbol{\lambda} - \boldsymbol{\lambda}^h\right) \ d\Gamma = \frac{h}{k} \int_{\Gamma_\mathrm{D}} \boldsymbol{\mu}^h \cdot \left(\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u}^h)\right) \ d\Gamma
\tag{34}
$$

In (16) the solution upon completion of the iterative procedure has been considered, so that the traction $\boldsymbol{T}$ is calculated for the same $\mathbf{u}^h$ as the energy. Operating as in the previous section, the multipliers can be eliminated to obtain the following orthogonality property of our formulation:

$$
a\left(\mathbf{u} - \mathbf{u}^h, \mathbf{v}^h\right) + \frac{k}{h} \int_{\Gamma_\mathrm{D}} \left(\mathbf{u} - \mathbf{u}^h\right) \cdot \mathbf{v}^h \ d\Gamma - \int_{\Gamma_\mathrm{D}} \left(\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u}^h)\right) \cdot \mathbf{v}^h \ d\Gamma = 0
\tag{35}
$$

Adding and subtracting the polynomial approximation of the traction in the Dirichlet boundary $\boldsymbol{T}(\mathbf{u})$ into the last integral of Equation 35 we obtain:

$$
Q(\mathbf{u} - \mathbf{u}^h, \mathbf{v}^h) - \int_{\Gamma_\mathrm{D}} \left(\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u})\right) \cdot \mathbf{v}^h \ d\Gamma = 0
\tag{36}
$$

Taking into account the stability of functional Q (Equation (31)) and using Equation (36) it follows that for any $\mathbf{w}^h \in \mathscr{U}^h$:

$$
\|\mathbf{u}^h - \mathbf{w}^h\|_{\mathscr{U}^h} \leq \frac{1}{\beta} \frac{Q(\mathbf{u}^h - \mathbf{w}^h, \mathbf{v}^h)}{\|\mathbf{v}^h\|_{\mathscr{U}^h}} = \frac{1}{\beta} \frac{Q(\mathbf{u}^h - \mathbf{u}, \mathbf{v}^h) + Q(\mathbf{u} - \mathbf{w}^h, \mathbf{v}^h)}{\|\mathbf{v}^h\|_{\mathscr{U}^h}} \leq
$$

22

$$\frac{1}{\beta} \frac{\left| \int_{\Gamma_{\mathrm{D}}} \left(\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u}) \right) \cdot \mathbf{v}^h \, d\Gamma \right| + Q(\mathbf{u} - \mathbf{w}^h, \mathbf{v}^h)}{\|\mathbf{v}^h\|_{\mathscr{U}^h}}$$

Now we use the Cauchy-Schwartz inequality and the continuity of functional Q with constant C to obtain:

$$\begin{aligned}
\|\mathbf{u}^h - \mathbf{w}^h\|_{\mathscr{U}^h} \leq & \frac{C}{\beta} \|\mathbf{u} - \mathbf{w}^h\|_{\mathscr{U}^h} + \frac{1}{\beta} \|\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u})\|_{L^2} \frac{\|\mathbf{v}^h\|_{L^2}}{\|\mathbf{v}^h\|_{\mathscr{U}^h}} \leq \\
& \frac{C}{\beta} \|\mathbf{u} - \mathbf{w}^h\|_{\mathscr{U}^h} + \frac{h^{1/2}}{\beta} \|\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u})\|_{L^2}
\end{aligned} \tag{37}$$

Finally we obtain the optimal convergence result:

Proposition 3. *Let $\mathbf{u}^h$ be the solution of problem* (18) *with $k > C_E C_p C_r$ ($\kappa > C_p C_r$) and $\mathbf{u}$ the exact solution. Then it holds that*

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathscr{U}^h} \leq \mathcal{O}(h^p) \tag{38}$$

p being the degree of the polynomial interpolation.

Proof: Let $\mathbf{w}^h$ be any function in the finite element space $\mathscr{U}^h$. We can write

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathscr{U}^h} = \|\mathbf{u} - \mathbf{w}^h - \mathbf{u}^h + \mathbf{w}^h\|_{\mathscr{U}^h} \leq \|\mathbf{u} - \mathbf{w}^h\|_{\mathscr{U}^h} + \|\mathbf{u}^h - \mathbf{w}^h\|_{\mathscr{U}^h}$$

Now we use Equation (37), the best approximation property of the finite element space (Equation 19) and property 4 of the smooth stress field $\boldsymbol{T}$ (Equation (23))

$$\|\mathbf{u} - \mathbf{u}^h\|_{\mathscr{U}^h} \leq \|\mathbf{u} - \mathbf{w}^h\|_{\mathscr{U}^h} + \frac{h^{1/2}}{\beta} \|\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{T}(\mathbf{u})\|_{L^2} \leq \mathcal{O}(h^p) + \mathcal{O}(h^{p+3/2})$$

7. Numerical examples

In this section we illustrate the capabilities of the proposed formulation and explore the limitations of the methods. Three numerical examples with exact solution were solved and used to check the convergence of the Richardson iterations, the convergence rate of the finite element solution as the mesh is refined and the effect of the stability constant k on convergence. They were also used to compare the proposed method with Nitsche's method showing that, in general, both methods have similar behavior. The third example

highlights the robustness of the proposed method in cases where Nitsche's method fails to provide the desired accuracy, *i.e.*, when the boundary of the domain comes close to the element faces. An additional example shows the behavior of the proposed method, without any modifications, when elasto-plastic behavior of the material is considered. This example is particularly interesting as, to the authors' knowledge, Nitsche's method has not been used to solve these type of problem because the formulation of the method required for plasticity has not been derived. Linear $(\mathcal{L}_8)$ and quadratic $(\mathcal{Q}_{20})$ elements are used in the examples.

7.1. Example 1: Tilted hexahedron

In the first example we consider an infinite domain subjected to 4th order polynomial displacements and plain strain conditions. The exact solution reads as:

$$E = 1000, \; \nu = 0.3$$

$$u_x = -\frac{25}{192} + \frac{75}{64}x^2 - \frac{25}{24}x^4 - \frac{25}{4}y+$$

$$\frac{25}{4}x^2y - \frac{25}{8}y^2 + \frac{25}{8}x^2y^2 \qquad \sigma_{xx} = \frac{5}{2}x - 3x^3 + 8xy + 4xy^2$$

$$u_y = \frac{20}{3}x + \frac{65}{12}x^3 + \frac{65}{12}x^3y - 10xy^2 - \frac{10}{3}xy^3 \qquad \sigma_{yy} = \frac{5}{8}x + \frac{14}{3}x^3 - 18xy - 9xy^2$$

$$u_z = 0$$

$$\sigma_{xy} = \frac{1}{6} + 9x^2 - \frac{5}{2}y + 9x^2y - 4y^2 - \frac{4}{3}y^3$$

$$t_{vx} = \frac{-E}{1+\nu}(1+y)$$

$$\sigma_{zz} = -0.3(\sigma_{xx} + \sigma_{yy})$$

$$t_{vy} = \frac{-E}{1+\nu}(1-x)$$

$$\sigma_{yz} = \sigma_{zx} = 0$$

$$t_{vz} = 0$$

A finite portion of the infinite domain defined by a tilted hexahedron was considered for the analysis. The known values of the displacements were imposed as Dirichlet boundary conditions on the entire external surfaces of the domain. Figure 5 shows the exact geometry of the problem embedded in the Level 3 Cartesian grid. The same figure also shows the surface triangulation used to evaluate numerically the contour integrals.

The problem was solved using a sequence of Cartesian meshes obtained by element subdivision starting from a Level 2 mesh having 64 hexahedral elements (Level 0 has a single element and Level 1 has eight elements).

In order to test the influence of the stability constant on the convergence of the proposed iterative method and on the discretization error, the same
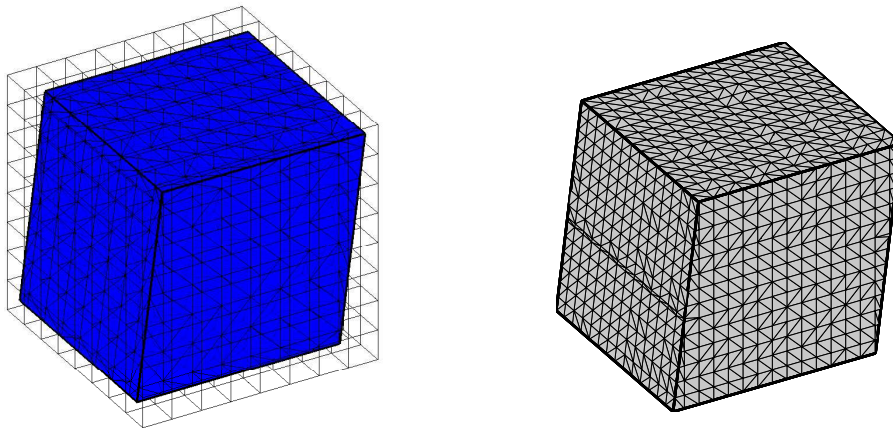
Figure 5: Example 1. Left: Geometry embedded in a Level 3 Cartesian mesh. Right: Triangulation of the surface in a Level 4 Cartesian mesh.

mesh was analyzed with κ ranging from 0.04 to $4 \cdot 10^6$. We also considered Nitsche's method in the convergence analysis for comparison purposes.

The convergence of the normalized residual (Equation (25)) is plotted in Figure 6 for different values of κ using linear and quadratic elements of the Level 2 Cartesian mesh. As predicted, small values of the penalty constant can cause very slow convergence or even loss of convergence in the iterative method. The greater the value of κ the faster the convergence, but very large values of the penalty constant can increase the discretization error of the finite element solution. To illustrate this the energy norm of the error is plotted in Figure 7 for different values of the penalty constant considering the Level 2 and 3 meshes (similar behavior was obtained for the rest of the mesh levels), and compared with the results obtained with Nitsche's method. It can be seen that the proposed technique gave a wide range of κ values, from 4 to $4 \cdot 10^3$, for which the level of the error remains essentially unaffected for both $\mathcal{L}_8$ and $\mathcal{Q}_{20}$ elements. However, the values of κ for which the error level remains unaffected is narrower in Nitsche's method, ranging from only $\kappa = 40$ to $\kappa = 4 \cdot 10^3$ for $\mathcal{L}_8$ elements and from $\kappa = 400$ to $\kappa = 4 \cdot 10^3$ for $\mathcal{Q}_{20}$ elements. It can also be observed that both techniques provide similar error levels for high levels of κ ($\kappa \geq 40$ for $\mathcal{L}_8$ elements and $\kappa \geq 400$ for $\mathcal{Q}_{20}$ elements).

In Figure 8 the energy and the L^2 norms of the discretization error are plotted as a function of the mesh size for $\mathcal{L}_8$ and $\mathcal{Q}_{20}$ elements, and for both the
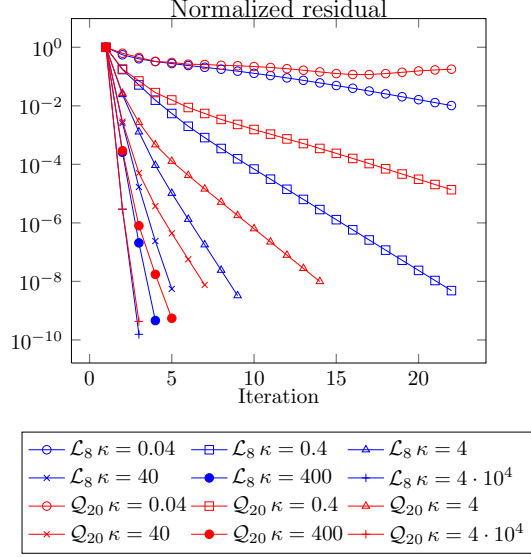
25

Figure 6: Example 1. Convergence of the residual in the Richardson iterations for the Level 2 mesh and different values of κ using linear and quadratic elements.
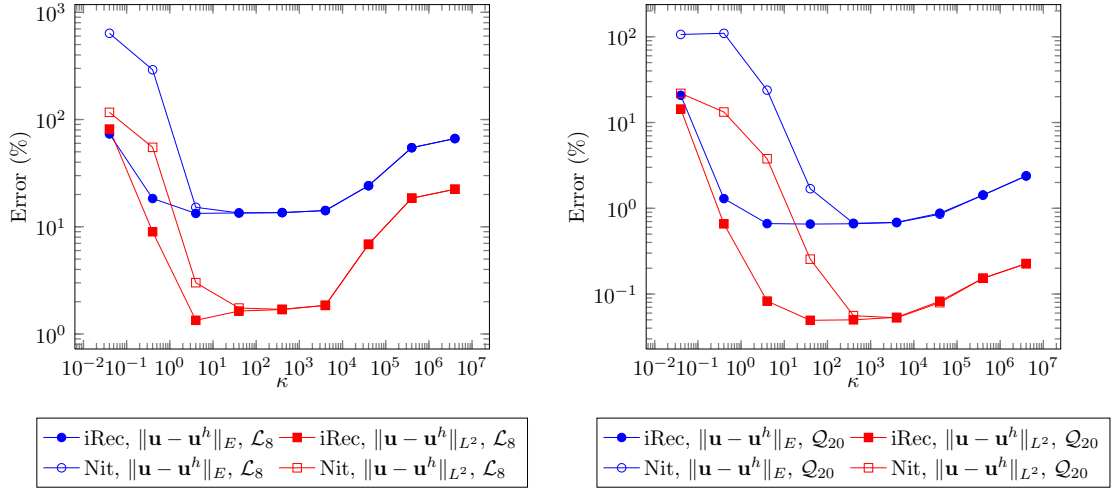

Figure 7: Example 1. Discretization error in energy norm and L^2-norm (%) for different values of the penalty constant κ. Results for Level 2 Cartesian mesh using linear and quadratic elements. The proposed method is denoted as iRec and Nitsche's method as Nit.

proposed technique (iRec) and Nitsche's method (Nit). The triangles in this figure show the theoretical optimal convergence rate that can be achieved. Three different values of κ were considered, taking into account the theoretical value $\kappa > C_p C_r$ that provides optimal convergence ($C_p = 13$ for $\mathcal{L}_8$ and $C_p = 21$ for $\mathcal{Q}_{20}$ elements, assuming that $C_r = 1$), $\kappa = 4$, $\kappa = 40$ and $\kappa = 400$.

As $\kappa = 4$ is lower than $C_p C_r$, the optimal convergence rate is not ensured. Indeed it is only achieved using the proposed method for $\mathcal{L}_8$ elements. The expected behavior in terms of convergence rate is observed in the numerical results for $\kappa = 40$ using the proposed method both for linear and quadratic elements, whereas Nitsche's method was only able to recover the optimum convergence rate for $\kappa = 400$ and linear elements. A reduction of the convergence rate for Nitsche's method can be seen in the last refinement step for $\mathcal{Q}_{20}$ elements. Nitsche's method would therefore require a κ higher than 400 to be able to successfully recover the optimum convergence rate during the entire refinement process for $\mathcal{Q}_{20}$ elements.

This example showed that, if a sufficiently high value of κ is used, the proposed technique provides results similar to those obtained with Nitsche's method. However, it is able to provide accurate results, with the optimal convergence rate, for considerably lower values of the stabilization parameter κ than Nitsche's method.

7.2. Example 2: Spherical domain

The second example uses a sphere as the exact geometry of the problem. Figure 9 shows the Level 3 and Level 4 meshes. The exact solution of the problem considered in this domain is a fourth-order polynomial defined in an infinite domain, which is imposed as a Dirichlet boundary condition on the surface of the sphere. The exact solution of the problem reads as:
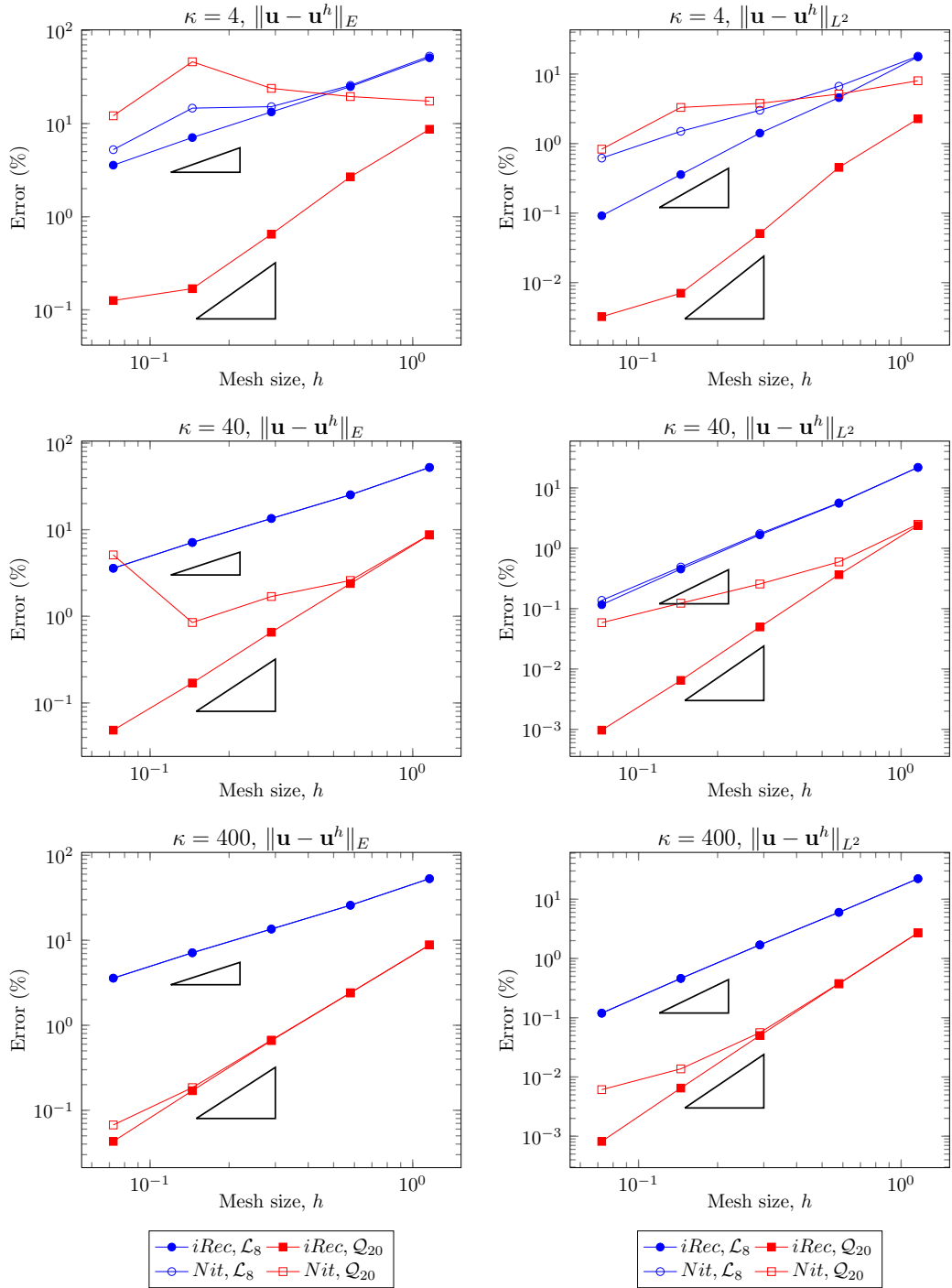
Figure 8: Example 1. Discretization error in energy norm as a function of the mesh size for $\mathcal{L}_8$ and $\mathcal{Q}_{20}$ elements. The triangles show the optimal convergence rate.

28

$$E = 1000, \; \nu = 0.3$$

$$A = \frac{E}{(2\,\nu - 1)\,(\nu + 1)}$$

$$u_x = x^4 + 3\,y\,x^3 - 2\,x\,z^2 + y\,x\,z$$

$$u_y = 7\,x^2\,y + y^4 - 2\,z\,y^3$$

$$u_z = -3\,x^2\,z^2 + 2\,y\,x\,z + z^3$$

$$b_x = A\left[(1 - \nu)\,\left(12\,x^2 + 18\,y\,x\right) + \nu\,\left(14\,x + 2\,y - 12\,x\,z\right) + \left(\nu - \frac{1}{2}\right)\,\left(-2\,y + 12\,x\,z - 10\,x\right)\right]$$

$$\sigma_{xx} = -A\left[\nu\,\left(7\,x^2 + 4\,y^3 - 6\,z\,y^2 - 6\,x^2\,z + 2\,y\,x + 3\,z^2\right) - (\nu - 1)\,\left(4\,x^3 + 9\,y\,x^2 - 2\,z^2 + y\,z\right)\right]$$

$$\sigma_{yy} = -A\left[\nu\,\left(-6\,x^2\,z + 2\,y\,x + 3\,z^2 + 4\,x^3 + 9\,y\,x^2 - 2\,z^2 + y\,z\right) - (\nu - 1)\,\left(7\,x^2 + 4\,y^3 - 6\,z\,y^2\right)\right]$$

$$\sigma_{zz} = -A\left[\nu\,\left(7\,x^2 + 4\,y^3 - 6\,z\,y^2 + 4\,x^3 + 9\,y\,x^2 - 2\,z^2 + y\,z\right) - (\nu - 1)\,\left(-6\,x^2\,z + 2\,y\,x + 3\,z^2\right)\right]$$

$$\sigma_{xy} = \frac{E}{2\,(1 + \nu)}\,\left(14\,x\,y + x\,z + 3\,x^3\right)$$

$$\sigma_{yz} = \frac{E}{2\,(1 + \nu)}\,\left(2\,x\,z - 2\,y^3\right)$$

$$\sigma_{xz} = \frac{E}{2\,(1 + \nu)}\,\left(x\,y - 4\,x\,z + 2\,y\,z - 6\,x\,z^2\right)$$

$$b_y = A\left[(\nu - 1)\,\left(12\,y\,z - 12\,y^2\right) - \left(\nu - \frac{1}{2}\right)\,\left(9\,x^2 + 14\,y + z - 2\,x\right) + \nu\,\left(9\,x^2 + 2\,x + z\right)\right]$$

$$b_z = A\left[\left(\nu - \frac{1}{2}\right)\,\left(6\,y^2 + 6\,z^2 + 4\,z - y\right) + \nu\,\left(y - 4\,z - 6\,y^2\right) - (\nu - 1)\,\left(6\,z - 6\,x^2\right)\right]$$
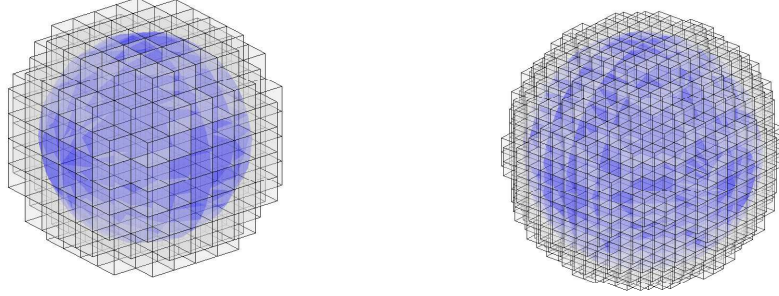

Figure 9: Spherical geometry embedded in a Cartesian mesh. Level 3 and Level 4 meshes.

In this example we use a technique that considers the exact geometry of the problem in the evaluation of volume and surface integrals. These integrals will be exactly evaluated up to the numerical integration errors [41]. Whatever the method used to impose Dirichlet boundary conditions in immersed boundary methods, considering the exact geometry of the domain in the evaluation of volume and surface integrals is necessary to achieve the optimal convergence rate for both linear and quadratic elements. This is
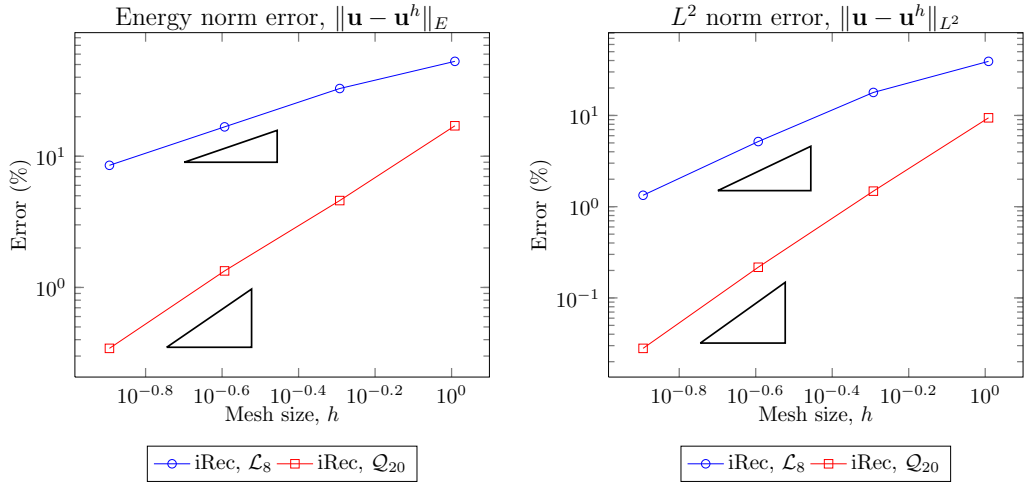
Figure 10: Example 2. Discretization error in energy norm for $\mathcal{L}_8$ and $\mathcal{Q}_{20}$ elements using $\kappa = 40$. The triangles show the optimal convergence rate.

because approximations to the actual geometry, for example by a faceted representation of the surface, could lead to geometrical modeling errors that could spoil the convergence rate of the numerical method. Figure 10 shows the discretization error in energy and L^2 norms. The optimal values of the error convergence rates, represented by the triangles shown in the graphs, are obtained in all cases. We recall that the theoretical rate of convergence for linear elements is 1 in energy norm and 2 in L^2-norm. For quadratic elements it is 2 in energy norm and 3 in L^2-norm. The exact slopes of the finite element solution for $\mathcal{L}_8$ elements are 0.69, 0.97 and 0.97 in energy norm and 1.13, 1.79 and 1.96 in L^2-norm, whilst for $\mathcal{Q}_{20}$ elements the values are 1.89, 1.78 and 1.96 in energy norm and 2.67, 2.77 and 2.95 in L^2-norm. The value of the penalty constant was $\kappa = 40$.

7.3. Example 3: Cubic domain parallel to the Cartesian grid

Let us consider the exact solution of the fourth-order polynomial given in the second example, in a hexahedral domain with faces parallel and equidistant to the faces of the embedding mesh. This solution was used to impose the Dirichlet boundary condition on the surface. Figure 11 shows two Cartesian meshes of Level 3 corresponding to two different bounding boxes that could be used to analyze this geometry. Let η represent the ratio of the volume of domain contained in the boundary elements to the total volume of these

elements ($\eta = 10\%$ in Figure 11 left and $\eta = 90\%$ in Figure 11 right). We will use η to represent the different relative positions of the surface in the intersected boundary elements.
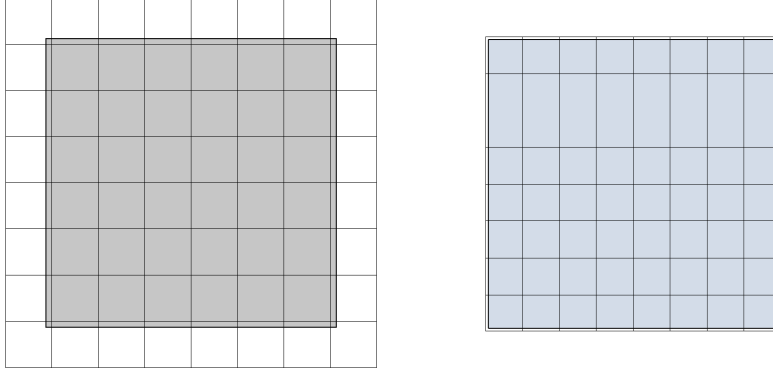

Figure 11: Example 3. Cartesian meshes with different values of the parameter η.

In order to test the influence of η on the error of the solution, sequences of uniformly refined meshes of $\mathcal{L}_8$ and $\mathcal{Q}_{20}$ elements were generated by adjusting the bounding box and the mesh level to keep the parameter η constant. The values of η selected for the analyses were 50%, 25%, 10%, 5%, 1%, 0.5%, 0.1% and 0.05%. This is a challenging problem for Nitsche's method, as the constant that provides stability C_N (Equation (14)) increases as η is reduced and cannot be bounded for $\eta \to 0$.

With the numerical analyses of this problem we try to show that the proposed method is able to provide accurate solutions even when Nitsche's method fails to do so, *i.e.* for small values of η, but is numerically equivalent to Nitsche's method for larger values of η. In fact we compared three techniques, denoted as $iRec$ - the iterative method proposed in this paper, $iNit$ - the iterative method described in Section 3.3 and Nit - the Nitsche's method.

The magnitude used to compare the results obtained from these three techniques is the exact L^2-norm error of the tractions on the Dirichlet boundaries $\left\| \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n} \right\|_{L^2,\Gamma_\mathrm{D}}$. Figure 12 shows in logarithmic scale the evolution of this magnitude for meshes of Level 3 and Level 4. The graphs for this example show that, in general terms, reducing η has a negative effect on the results obtained by all three techniques, but especially in the case of Nitsche's method. For $\eta = 0.05\%$, the best results obtained with Nitsche's method, which were for the highest value of the stabilization parameter ($\kappa = 4 \cdot 10^3$),

are unacceptable ($\left\|\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}\right\|_{L^2,\Gamma_{\mathrm{D}}} = 46258.6\%$ for the Level 3 mesh and $\left\|\boldsymbol{\sigma}(\mathbf{u})\mathbf{n} - \boldsymbol{\sigma}(\mathbf{u}^h)\mathbf{n}\right\|_{L^2,\Gamma_{\mathrm{D}}} = 7786.7\%$ for Level 4). Reducing η leads to increasingly higher required values of κ. On the one hand the use of these high κ values gives a high weight to the satisfaction of the Dirichlet boundary conditions and leads to an improvement of the magnitude considered in the comparisons in this example. However, on the other hand it increases the condition number of the system matrix, and might overweight the imposition of the Dirichlet BC on the surface at the expense of reducing the weight of the terms that account for the energy in the volume (similar ideas have already been put fordward by [19]). This implies high error levels in the results obtained in the elements cut by the Dirichlet boundary. This problem would be critical if the results obtained in these elements were magnitudes of interest to the analyst.

The $iNit$ curves show a slightly better performance than the Nit curves for $\kappa = 4 \cdot 10^3$, and more significantly for $\kappa = 400$. However, further reducing the value of κ prevents the convergence of the iterative process in the full range of values of the volume ratio η (convergence was only obtained when $\kappa = 4$ for $\eta \geq 10\%$ and, when $\kappa = 40$, for $\eta \geq 1\%$ for the Level 3 mesh and for $\eta \geq 0.5\%$ in the case of Level 4). The results given in figure show that the best results are obtained with the proposed method ($iRec$ curves). These curves are similar to the Nit and $iNit$ curves for $\kappa = 4 \cdot 10^3$, but show a considerable improvement when κ is reduced. The optimum performance shown in the graphs for the $iRec$ curves is at $\kappa = 4$ with the highest error levels around 30% and 26% (only obtained for $\eta \leq 0.1\%$) respectively for the Level 3 and Level 4 meshes, $i.e.$ several orders of magnitude smaller than with the other two methods.

7.4. Example 4: Plasticity

In this last example we check the performance of the proposed technique when used to analyze a mechanical component considering elasto-plastic behavior of the material. The component analyzed is a rectangular plate with a central hole subjected to uniaxial monotonic traction, as represented in Figure 13. The highlighted 1/8 of the plate with the appropriate symmetry boundary conditions was used in the analyses. The material behavior is a von Mises plasticity bilinear model with yield stress $S_y = 24$ units of pressure and slope of the plastic zone $H = 225$ units of pressure. The Young modulus is $E = 1000$ units of pressure and the Poisson's ratio $\nu = 0.3$. The maximum
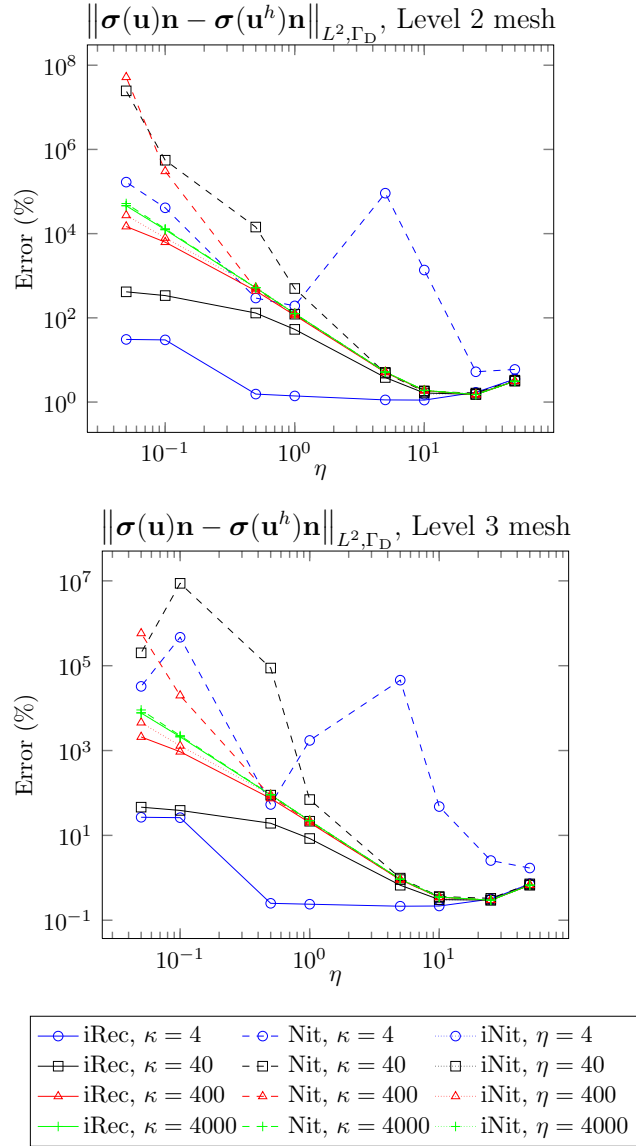
Figure 12: Example 3. Fourth-order polynomial in a cube. L^2-error of the traction field on the Dirichlet boundary for Level 2 and Level 3 meshes.
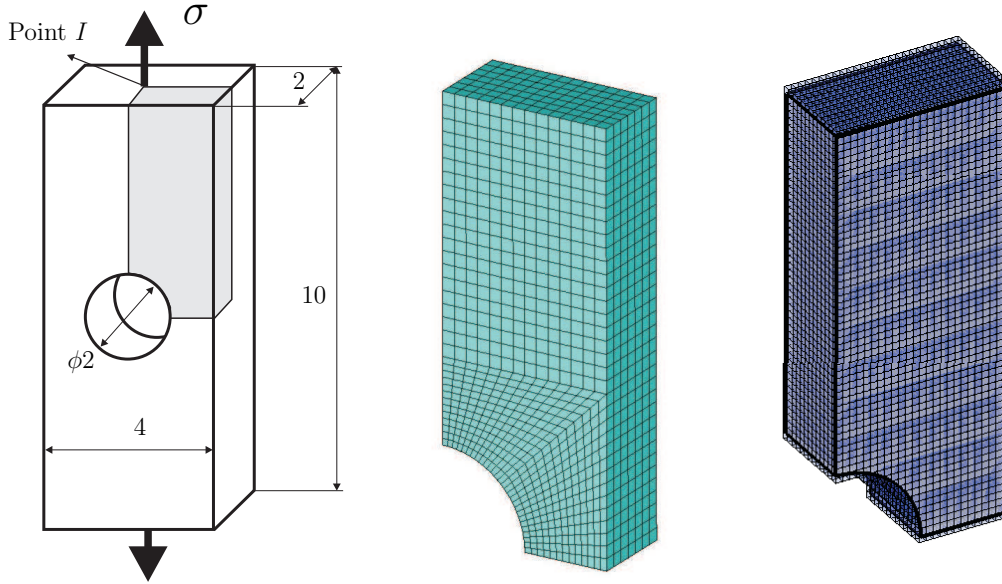
Figure 13: Example 4. Plate under traction. Geometrical model and finite element models. Centre: mesh of the reference model. Right: Level 5 Cartesian grid mesh.

traction applied is $\sigma_{max} = 0.9 \cdot S_y$. We use the proposed method to apply the Dirichlet boundary conditions on the symmetry boundaries. Although the theoretical analysis done in the paper to obtain the value of k is only valid for linear elasticity, we use the same values for non-linear problems, such as, the present example. Therefore, the value of the stabilizing constant is $\kappa = 40$ ($k = 100$ using the elastic material properties).

The reference solution was evaluated using ANSYS®[42]. We compared the displacement of Point **I** (shown in Figure 13) as a function of the load and the final distribution of the plastic zone, using the reference and the Cartesian grid methods. A sequence of analyses using mesh levels 2, 3, 4 and 5, with element faces not coplanar to the symmetry surfaces, was run to test the convergence of the solutions obtained by the Cartesian grid method to the reference. The size of the elements in the reference model was similar to that of the Level 5 mesh, although the Ansys model gave smaller element sizes around the hole. The number of degrees of freedom of the Cartesian grid models was 396, 1935, 10392 and 65058, whereas the reference model had 92904 degrees of freedom. A comparison between the reference mesh and the Cartesian grid for Level 5 is shown in Figure 13.
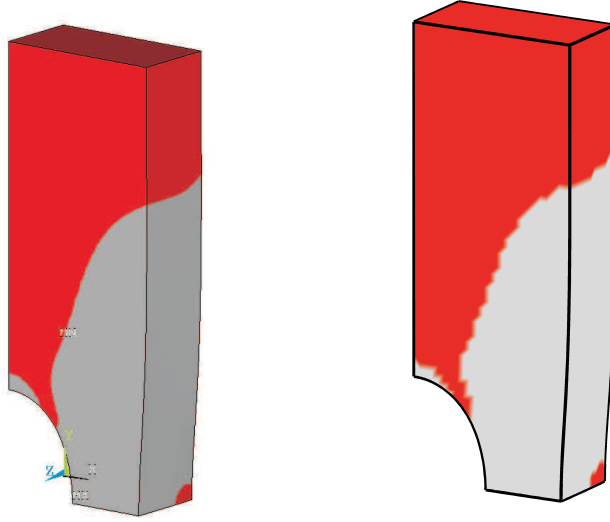
34

Figure 14: Example 4. Comparison of the FE models. Left: mesh of the reference model (Ansys). Right: Level 5 Cartesian grid mesh.

Figure 14 compares the zone of the models that underwent plasticity (shown in gray) evaluated by Ansys and by our Cartesian grid implementation. It can be observed that both zones are similar, even though different graphical representation techniques are used by each of the codes.

Figure 15 shows the load-displacement curves obtained from the different analyses. It can be clearly observed that the curves obtained with the Cartesian grid smoothly converge to the reference curve, thus showing that the technique of imposing Dirichlet boundary conditions proposed in this paper can be used to provide similar results to those obtained with standard finite element implementations.

8. Conclusions

This paper describes a novel method of imposing Dirichlet boundary conditions suitable for immersed boundary Cartesian meshes in an approach based on the stabilized Lagrange multipliers method. This approach allows the Lagrange multipliers to be condensed element-by-element during the assembly process. The stabilization term is evaluated by using a smoothed stress field obtained from the Superconvergent Patch Recovery technique. An iterative
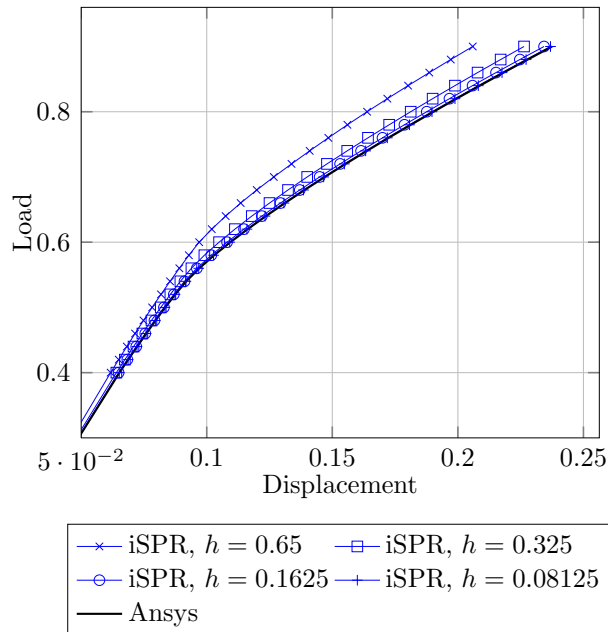
35

Figure 15: Example 4. Load-displacement curve: Convergence of the Cartesian grid solution to the Ansys solution with mesh refinement.

procedure is defined to update the stabilizing term and the global convergence of the procedure is proved for a sufficiently large value of the stabilizing parameter κ. The numerical examples show the influence of κ both on the finite element error and on the convergence of the Richardson iterations. In the examples given, the finite element solution remains unchanged for a wide range of values of κ ($\kappa \in [40 - 4 \cdot 10^3]$) and the convergence of the Richardson iteration is verified for $\kappa > 4$.

The numerical results show that the optimal convergence rate of the finite element solution is obtained for both linear and quadratic elements, provided that the geometry of the problem is accurate enough. In particular, if the geometry is approximated using a linear subtriangulation the solution using linear element has an optimal convergence rate whilst the convergence using quadratic elements is suboptimal.

The numerical comparisons with Nitsche's method showed that in the general case, and especially for high values of the penalty parameter of the stabilization term κ, both techniques provide similar results. However, the proposed technique proved to be robust in providing accurate results even

when the boundary of the domain comes very close to the element faces, where Nitsche's method is not robust. Moreover, the results obtained with the proposed technique improve as κ is reduced consequently with less likelihood of obtaining high condition numbers.

The proposed technique with no modifications was successfully applied to analyzing a problem with non-linear material behavior. Despite the fact that the plasticity covered a large area of the Dirichlet boundary, applying the proposed technique provided similar results to those obtained by the standard finite element method. Although further improvements could be gained by adapting the recovery technique to account for plastic deformations, the preliminary results of this work open up the possibility of obtaining the benefits of embedded domain methods in this type of non-linear problem.

Acknowledgements

Bibliography

[1] J. A. Cottrell, T. J. R. Hughes, Y. Bazilevs, Isogeometric Analysis: Toward Integration of CAD and FEA, Wiley, 1st edition, 2009.

[2] S. Bordas, T. Rabczuk, J. Rodenas, P. Kerfriden, M. Moumnassi, S. Belouettar, Recent advances towards reducing the meshing and re-meshing burden in computational sciences, Computational Technology Reviews 2 (2010) 51–82.

[3] T. Strouboulis, K. Copps, I. Babuška, The generalized finite element method: an example of its implementation and illustration of its performance, International Journal for Numerical Methods in Engineering 47 (2000) 1401–1417.

[4] M. Moumnassi, S. Belouettar, E. Bechet, S.P.A. Bordas, D. Quoirin, M. Potier-Ferry, Finite element analysis on implicitly defined domains:

An accurate representation based on arbitrary parametric surfaces., Computer Methods in Applied Mechanics and Engineering 200 (2011) 774–796.

[5] E. Nadal, J. J. Ródenas, J. Albelda, M. Tur, J. E. Tarancón, F. J. Fuenmayor, Efficient Finite Element Methodology based on Cartesian grids: Application to structural shape optimization, Abstract and Applied Analysis (2013) 1–19.

[6] G. Xu, E. Atroshchenko, W. Ma, E. Bordas, Geometry-Independent Field approximation: CAD-Analysis Integration, geometrical exactness and adaptivity., Computer Methods in Applied Mechanics and Engineering (2014) .

[7] E. Nadal, Cartesian grid FEM (cgFEM): High performance h-adaptive FE analysis with efficient error control. Application to structural shape optimization, Ph.D. thesis, Universidad Politécnica de Valencia, 2013.

[8] R. Sevilla, S. Fernández-Méndez, A. Huerta, 3D Nurbs-enhanced Finite Element Method (NEFEM), International Journal for Numerical Methods in Engineering 88 (2011) 103–125.

[9] F. Brezzi, M. Fortin, Mixed and hybrid finite element methods, Springer-Verlag New York, Inc., New York, NY, USA, 1991.

[10] E. Béchet, N. Moës, B. Wohlmuth, A stable Lagrange multiplier space for stiff interface conditions within the extended finite element method, International Journal for Numerical Methods in Engineering 78 (2009) 931–954.

[11] M. Hautefeuille, C. Annavarapu, J. Dolbow, Robust imposition of Dirichlet boundary conditions on embedded surfaces, International Journal for Numerical Methods in Engineering 90 (2012) 40–64.

[12] H. J. Barbosa, T. J. Hughes, The finite element method with Lagrange multipliers on the boundary: circumventing the Babuska-Brezzi condition, Computer Methods in Applied Mechanics and Engineering 85 (1991) 109-128.

[13] H. J. Barbosa, T. J. Hughes, Boundary Lagrange multipliers in Finite Element Methods: Error analysis in natural norms, Numerische Mathematik 62 (1992) 1–15.

[14] A. Hansbo, P. Hansbo, An unfitted Finite Element Method based on Nitsche method for elliptic interface problems, Computer Methods in Applied Mechanics and Engineering 191 (2002) 5537 – 5552.

[15] P. Hansbo, C. Lovadina, I. Perugia, G. Sangalli, A Lagrange multiplier method for the finite element solution of elliptic interface problems using non-matching meshes, Numerische Methematik 100 (2005) 91 – 115.

[16] J.E. Dolbow, I. Harari, An efficient finite element method for embedded interface problems, International Journal for Numerical Methods in Engineering 78 (2009) 229–252.

[17] J. D. Sanders, J. E. Dolbow, T. A. Laursen, On methods for stabilizing constraints over enriched interfaces in elasticity, International Journal for Numerical Methods in Engineering 78 (2009) 1009–1036.

[18] J. Haslinger, Y. Renard, A new fictitious domain approach inspired by the extended Finite Element Method, SIAM J. Numer. Anal. 47 (2009) 1474–1499.

[19] B. Schott, W. Wall, A new face-oriented stabilized XFEM approach for 2D and 3D incompressible Navier-Stokes equations, Computer Methods in Applied Mechanics and Engineering 276 (2014) 233 – 265.

[20] E. Burman, P. Hansbo, Fictitious domain Finite Element methods using cut elements: Ii. a stabilized Nitsche method, Applied Numerical Mathematics 62 (2012) 328 – 341.

[21] S. Amdouni, P. Hild, V. Lleras, M. Moakher, Y. Renard, A stabilized Lagrange multiplier method for the enriched finite-element approximation of contact problems of cracked elastic bodies, ESAIM: Mathematical Modelling and Numerical Analysis 46 (2012) 813–839.

[22] C. Annavarapu, M. Hautefeuille, J. Dolbow, A robust Nitche's formulation for interface problems, Computer Methods in Applied Mechanics and Engineering 228 (2012) 44–54.

[23] J. D. Sanders, T. A. Laursen , J. E. Dolbow, A Nitsche embedded mesh method, Computational Mechanics 49 (2012) 243–257.

[24] C. Annavarapu, M. Hautefeuille, J. Dolbow, A Nitsche stabilized finite element method for frictional sliding on embedded interfaces. Part I: single interface., Computer Methods in Applied Mechanics and Engineering 268 (2014) 417–436.

[25] R. Codina, J. Baiges, Approximate imposition of boundary conditions in immersed boundary methods, International journal for numerical methods in engineering 80 (2009) 1379–1405.

[26] E. Burman, Projection stabilisation of Lagrange multipliers for the imposition of constraints on interfaces and boundaries (2012).

[27] M. Tur, J. Albelda, E. Nadal, J. J. Rodenas, Imposing Dirichlet boundary conditions in hierarchical Cartesian meshes by means of stabilized Lagrange multipliers, International Journal for Numerical Methods in Engineering 98 (2014) 399–417.

[28] A. Lew, G. Buscaglia, A discontinuous-Galerkin-based immersed boundary method, International journal for numerical methods in engineering 76 (2008) 427–454.

[29] F. Liu, R. I. Borja, Stabilized low-order finite elements for frictional contact with the extended Finite Element Method, Computer Methods in Applied Mechanics and Engineering 199 (2010) 2456 – 2471.

[30] P. B. Bochev, C. R. Dohrmann, M. D. Gunzburger, Stabilization of low-order mixed finite elements for the Stokes equations, SIAM J. Numer. Anal. 44 (2006) 82–101.

[31] G. Barrenechea, F. Chouly, A local projection stabilized method for fictitious domains, Applied Mathematics Letters 25 (2012) 2071–2076.

[32] J. Baiges, R. Codina, F. Henke, S. Shahmiri, W. A. Wall, A symmetric method for weakly imposing Dirichlet boundary conditions in embedded Finite Element meshes, International Journal for Numerical Methods in Engineering 90 (2012) 636–658.

[33] O. C. Zienkiewicz, J. Z. Zhu, The Superconvergent Patch Recovery and a posteriori error estimates. Part 2: Error estimates and adaptivity, International Journal for Numerical Methods in Engineering 33 (1992) 1365–1382.

[34] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique 8 (1974) 129–151.

[35] J. Pitkaranta, A conforming Finite Element Method with Lagrange multipliers for the biharmonic problem, ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique 14 (1980) 309–324.

[36] R. Stenberg, On some techniques for approximating boundary conditions in the Finite Element Method, Journal of Computational and Applied Mathematics 63 (1995) 139 – 148.

[37] W. Jiang, C. Annavarapu, J. Dolbow, I. Harari, A robust Nitsche's formulation for interface problems with spline-based finite elements., International Journal for Numerical Methods in Engineering (2015) .

[38] A. Gerstenberger, W. A. Wall, An embedded Dirichlet formulation for 3D continua, International Journal for Numerical Methods in Engineering 82 (2010) 537–563.

[39] J. J. Ródenas, M. Tur, F. J. Fuenmayor, A. Vercher, Improvement of the Superconvergent Patch Recovery technique by the use of constraint equations: the SPR-C technique, International Journal for Numerical Methods in Engineering 70 (2007) 705–727.

[40] Y. Saad, Iterative methods for sparse linear systems, Applied Mathematical Sciences, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2003.

[41] O. Marco, R. Sevilla, Y. Zhang, J. J. Rodenas, M. Tur, Efficient integration of exact geometries by means of 3D Cartesian grids, Accepted DOI:10.1002/nme, International Journal for Numerical Methods in Engineering (2015).

[42] Ansys 15.0 ANSYS Academic Research, Release 15.0, Help System, AN-SYS, Inc., Philadelphia, PA, USA.